

# 博士論文

報酬ベース自律エージェントにおける間接的行動の設計

公立はこだて未来大学大学院 システム情報科学研究科

システム情報科学専攻

大 宮 健 太

2010 年 02 月

# Doctoral Thesis

Design of Indirect Behavior for Reward Based Autonomous Agents

By

Kenta Oomiya

Graduate School of Information Science

Future University Hakodate

February, 2010

## **Abstract:**

Some living creatures not only have behavior for achieving its purpose directly but also indirect behavior supporting direct behavior. Spiders evolved from beings with direct predation to beings with indirect predation by using their webs as traps. Thus, it is said that weaving webs can support the effectiveness of the direct predation. Indirect behavior such as weaving webs or structural objects provides probabilities as follows: (1) probabilities for preying on bugs which are physically stronger or bigger or faster than predators or (2) probabilities for improving the effectiveness of its predation. That is, spiders which do not weave webs can capture only bugs which are smaller and slower than them. In contrast, spiders which weave webs can capture flying bugs and bugs which are faster than them with their webs. In this paper, we regarded behavior achieving a purpose directly as “direct behavior” and indirect behavior which can improve effectiveness of direct actions as “meta-behavior”. We focused on the meta-behavior and studied about it. In natural environments, there are many examples of meta-behavior as the spiders’ example. These examples indicate that it is said that there are some difficulties introducing meta-behavior to an environment. When meta-behavior is introduced to an environment, it could be considered that rewards from direct behavior were decreased temporarily because agents have to select behavior except direct behavior. Therefore, to acquire meta-behavior stably and continuously could be difficult. There are many studies about acquisition of complicated behavior which consists of

simple actions, while there are not many studies about acquisition of such as meta-behavior. Therefore, we have tried to propose an approach to introduce meta-behavior for solving a problem. In order to study the approach, we applied it to a problem in game theory because of its theoretical and stable nature and then we examined the effectiveness of our approach. Furthermore we studied how to acquire appropriate meta-behavior in more complicated and unstable environments. We applied our approach to acquisition of constructive behavior of a nest for predation in a 2D environment and a 3D environment. As a result, we could examine the effectiveness of our approach introducing meta-behavior and propose autonomously acquisition of meta-behavior.

**Keywords:**

Meta-behavior, Autonomous Acquisition of Constructive Behavior, Virtual Creatures

## 概要：

自然の生物の行動の中には、ある目的を直接的に達成しようとする行動だけではなく、そのような行動をサポートするような行動が存在している。例えばクモ類の場合、直接追いかけて捕獲を行う種から、巣のような罠を作って獲物を捕獲する種に進化している。この直接的な捕獲に対するサポート行動が罠の構築行動であると言える。罠を作るという間接的行動を行う事によって、自身の身体能力を超える獲物を捕獲可能に出来たり、直接捕獲の場合よりもより多くの獲物を捕まえる事が出来たりするようになっている。クモ類の場合、巣を使わない種では、自分より小さく、移動速度の遅い獲物しか捕獲できないが、巣を使う種では、獲物を足止めする事で、空を飛ぶ獲物や自分より大きく速い獲物を捕獲し、捕食する事が出来る。本研究では、このような「ある目的を直接的に達成しようとする行動」を直接的行動、「直接的行動の効率や達成可能な事柄を増大させる事ができる間接的行動」をメタ行動と呼び、焦点を当てて研究を行う。メタ行動を導入する事で、効率性や達成可能な事柄を拡大できる例は自然界には多くあるが、その導入には難しい性質が存在していると考えられる。それは、メタ行動の導入時には、直接的行動以外の行動を取らざるを得ず、その分だけ利得が一時的に低下する事が多く、安定的・継続的にメタ行動の獲得が行われにくいという点である。捕獲行動の獲得の研究事例では、行動を組み合わせて複雑な行動を獲得する研究については数が多いが、いわゆるメタ行動の獲得に関する研究は少ない。そこで、本研究では、メタ行動の導入による問題解決アプローチの提案を目的として研究を行う。その為に、ゲーム理論のような論理的かつ不確定性の少ない問題への適用を通じて、メタ行動の導入の有効性の検証と獲得に関する実験を行った。その後、実用的な問題への適用を

行う為に必要となる、不確実性の高い環境におけるメタ行動の進化的獲得に関する実験を行った。獲物捕獲用の罠の構築行動獲得の実験を行い、結果として、メタ行動の導入による問題解決アプローチの有効性を確認し、メタ行動の獲得手法についても提案する事ができた。

**キーワード：**

メタ行動、造作物構築行動の獲得、仮想生物、

1. 序論.....	1
2. 関連研究.....	5
2.1 Virtual Creatures : Karl Sims の研究とその後続について .....	5
2.2 NeuroEvolution of Augmenting Topology (NEAT) .....	6
2.3 Evolutionary Robotics .....	7
2.4 Evolutionary Design.....	7
2.5 2次元上でのクモの巣の構築行動に関する研究 .....	11
2.6 社会的ジレンマに対するゲーム理論とエージェントアプローチ .....	11
2.7 生物学での関連研究.....	13
3. メタ行動の導入の枠組み .....	14
3.1 メタ行動の導入の枠組みのモデル .....	14
3.2 メタ行動の導入の枠組みを用いた各題材の表現.....	17
4. ゲームエージェントにおけるメタ行動 .....	19
4.1 共有地の悲劇 .....	19
4.1.1 概要.....	19
4.1.2 共有地の悲劇の定式化 .....	20
4.1.3 メタエージェントによる課税戦略 .....	24
4.2 メタエージェント化機能の導入.....	27
4.3 遺伝的アルゴリズムにおける評価関数について.....	28
4.4 拡張版の共有地の悲劇のシミュレーションについて .....	29
4.5 メタ行動導入の枠組みでの表現.....	32
4.6 シミュレーション実験とその結果について.....	32
4.7 拡張版共有地の悲劇におけるメタ行動の獲得について .....	35
4.8 4章についてのまとめ.....	35

5.	2次元シミュレーションにおける仮想生物.....	36
5.1	問題の性質について.....	36
5.2	自然生物におけるメタ行動.....	37
5.3	クモの生態について.....	38
5.3.1	身体的特徴について.....	38
5.3.2	造網行動について.....	40
5.3.3	補虫行動について.....	43
5.4	シミュレーションの概要.....	44
5.5	造作物構築行動の獲得のメタ行動の導入の枠組みによる表現.....	45
5.6	実験1 造作物構築行動の進化的獲得.....	46
5.5.1	獲物エージェント.....	47
5.5.2	捕獲者エージェント.....	49
5.5.2.1.	センシング.....	49
5.5.2.2.	基本的な行動.....	50
5.5.2.3.	意思決定.....	50
5.5.3	評価関数.....	52
5.5.4	実験1の設定.....	53
5.5.5	実験結果1.....	54
5.5.5.1.	評価関数1を用いた場合の実験.....	54
5.5.5.2.	評価関数2を用いた場合の実験.....	56
5.5.5.3.	センシングの仕方が異なる場合について.....	58
5.5.6	実験1まとめ.....	59
5.6	実験2 直接捕獲と間接捕獲の両方を獲得できる捕獲者による実験.....	60
5.6.1	環境設定の変更.....	61



5.6.2	直接捕獲用のセンサ .....	63
5.6.3	評価関数 .....	64
5.6.4	実験設定 .....	65
5.6.5	実験結果 1 獲物にブロック回避行動がない場合 .....	67
5.6.5.1.	直接捕獲と間接捕獲を別々に行った結果について .....	67
5.6.5.2.	直接捕獲と間接捕獲のどちらも選択できる場合 .....	68
5.6.6	実験結果 2 獲物がブロック回避行動を持つ場合 .....	71
5.6.6.1.	直接捕獲と間接捕獲を別々に行った場合の結果 .....	71
5.6.6.2.	直接捕獲と間接捕獲のどちらも選択できる場合 .....	72
5.6.7	直接捕獲と間接捕獲の選択 .....	75
5.7	第 5 章まとめ .....	77
6.	3次元物理シミュレーションにおける仮想生物 .....	78
6.1	シミュレーションの設定 .....	78
6.1.1	シミュレーション環境 .....	78
6.1.2	獲物エージェント .....	79
6.1.3	捕獲者エージェント .....	80
6.1.4	意思決定機構 .....	81
6.1.5	評価関数 .....	82
6.2	実験 1 構造物構築行動の進化的な獲得 .....	83
6.2.1	実験パラメータ .....	83
6.2.2	実験結果 .....	84
6.3	実験 2 中間評価の影響 .....	86
6.3.1	中間評価 .....	87
6.3.2	意思決定機構への変更 .....	87

6.3.3 実験パラメータ .....	88
6.4 第6章のまとめ .....	98
7. 結論 .....	99
7.1 結論 .....	99
7.2 メタ行動の設計論に向けて .....	100
7.3 今後の展望と予想される応用 .....	101
参考文献 .....	103
業績一覧 .....	109
査読付き論文 .....	109
国際発表 .....	109
国内学会・シンポジウム等における口頭発表 .....	111
国内学会・シンポジウムなどにおけるポスター発表 .....	112
謝辞 .....	113

## 1. 序論

科学の発展に伴い、生命に関する様々な謎が解明されてきているが、生物の脳や知能についての謎は未だ完全に明かされてはいない。今までの科学は、対象を分解して要素を取り出し、その要素の性質を調べる事で、対象の全体像を明らかにしてきた。しかし、生命や知能のようなシステムは、取り出した要素の性質を調べても、システム全体の挙動や性質の解明につながらない。それは、それらのシステム全体の挙動・性質が要素間の関係にも関連しているからである。特に生命系のシステムでは、要素を取り出してしまうと要素間の関係については調べる事ができない為、研究者の中には還元的アプローチでは限界があると考え人達もいる。近年では、このようないわゆる複雑系と呼ばれるシステムに対して、ボトムアップ的、構成論的アプローチが適用され、徐々にその性質が解明されてきている。構成論的アプローチでは、構成要素をモデル化し、複数の要素の振る舞いをシミュレーション上で再現し、要素間の相互作用を発生させ、システム全体としての振る舞いの創発を目指している。しかし、どのようにすれば創発が発生するのか、という事については未だ十分には体系化されておらず、個々の事例について個別の手法を適用している段階である。その理由としては、複雑系としての特徴を持つシステムは生物、脳、社会、交通など多岐にわたり、それぞれが個別の研究として大きな粒度を持っており、モデル化の仕方にも大きな違いがあることが 1 つの原因であると考えられる。このような状況を打破する為には、どのようにすれば創発現象が発生するのかという包括的な議論、モデルが必要になってくると考えられる。しかし、現状では、個別の事例に対する研究についても進展は十分ではないように感じる。即ち、現状では、個別の事例に対する研究を進め、事例を充実させる事が必要である

と考えられる。実際、多くの事例について盛んに研究が行われているが、数ある創発に関する研究領域の中でも、本研究では昆虫やクモなどにおける行動獲得に焦点を当てて研究を進める。人工知能・人工生命の分野で行動獲得に関する有名な研究としては、Rodney Brooks の Subsumption Architecture [37, 38] や Karl Sims の Virtual Creatures[1, 2]が挙げられる。これらの研究では、ある目的を達成する為の行動の獲得について研究を行っている。その一方で、ある目的を直接的に達成しようとする行動だけではなく、そのような行動をサポートするような行動についての研究事例はあまり多くないが、自然の生物の行動の中にはそのような間接的行動が存在している。

本論文でのメタ行動の定義は、他の行動をサポートするような間接的行動、である。これは、即ち、(1) ある行動の効率を向上させる事ができる間接的行動や、(2) ある行動だけでは問題解決できないような状況をその間接的行動を導入する事で解決可能であるような行動の事である。

クモ類の例を挙げて説明すると、直接追いかけて獲物を捕まえる行動に対する間接的行動として、捕獲用の罟を作る行動が挙げられる。クモ類は、造網行動を行わずに直接的に獲物を捕獲する種族と、造網行動によって巣を構築し、その巣を活用して獲物を捕獲する種族がいる。進化系統的には、直接捕獲を行う種族から、造網行動を行う種族が進化的に発生したと言われている。造網行動は、クモの身体的能力を補う事ができ、直接捕獲では捕獲できないようなクモよりも身体能力の高い獲物や、空を飛ぶ獲物を捕まえる事を可能とする。また、直接捕獲よりも効率的に獲物を捕まえる事が可能である。しかし、直接捕獲に比べて巣の構築コストがかかるうえに、一時的に獲物を捕獲できない期間が発生するなど、獲得に関して不利な条件が多く、造網行動の獲得には難しい。

上記の例で示した通り、安定的な獲得が難しい場合が多いが、その導入によ

って問題解決が見込める場合がある。しかし、メタ行動に関する研究はあまり多くなく、まだまだ研究の余地が残されている。

本研究では、メタ行動の導入による問題解決に焦点をあて、メタ行動を獲得する為の条件やメタ行動の獲得の方法について議論を行う。その為に、2種類の問題に対してメタ行動の導入を行い、段階的に研究を行った。1つ目の題材は、ゲーム理論の問題のひとつである共有地の悲劇である。この問題は、論理的かつ確定的なルールで表現されており、環境の挙動が予測可能な事例である。その様な環境において、メタ行動の導入によって全体の利得が上昇可能かどうか、実験を行った。2つ目の題材は、環境の挙動が予測しにくく、ランダム性の強い問題、即ち、共有地の悲劇よりも複雑な問題である捕獲用造作物の構築行動の獲得である。人工生命の領域で議論される、実環境を模した環境における捕獲者の行動獲得シミュレーションを用いて、2次元離散環境と3次元連続環境において実験を行った。2次元離散環境では、獲物を捕獲する為の造作物を構築するメタ行動と、獲物を直接捕獲する直接行動、その両方が発現しうる捕獲者による実験であり、人工ニューラルネットワークと遺伝的アルゴリズムの組み合わせ手法である **NeuroEvolution** 手法を用いて、どのような環境条件でメタ行動が有効に発現しうるのか、実験と考察を行い、コストとの関連でメタ行動が有効に発現する状況について実証的に明らかにした。3次元連続環境では、連続空間においてメタ行動の獲得が可能かどうか実験を行った。また、造作物の構築途中における構造的な評価が、造作物構築行動に与える影響について実験を行った。結果として、メタ行動の進化的な獲得について、遺伝的アルゴリズムや **NeuroEvolution** 手法による枠組みを示すことができた。また、直接的行動だけの環境よりもメタ行動を導入した環境の方が良いパフォーマンスを得る場合があることを示すことができた。そして、造作物構築途中での構造的な評価を行

い、意思決定の入力とすることで、行動パターンの切り替えを含むプロセスの獲得の可能性について示唆することができた。

本論文の構成を以下に述べる。第 2 章では、本研究の関連研究について述べる。第 3 章では、メタ行動の導入の枠組みについて詳細を述べ、続く第 4 章では、共有地の悲劇へのメタ行動導入事例について詳細を述べる。第 5 章以降では、造作物構築行動の獲得問題に対してメタ行動の導入を行う事例について詳細を述べる。第 5 章では、2 次元環境における造作物構築行動の獲得実験を、第 6 章では、3 次元環境における造作物構築行動の獲得実験について述べている。第 7 章では、研究の応用可能性について言及するとともに、本研究の総括を行った。

## 2. 関連研究

本章では、本研究で対象としている領域の関連研究について、本研究との関連を中心に言及を行う。関連研究として、大まかに以下の研究が挙げられる。

### 2.1 Virtual Creatures : Karl Sims の研究とその後続について

人工生命分野では、人工生命が環境に適応する為にその身体や行動を進化的・学習的に獲得する研究が数多くある。その先駆けとも言える研究は 1994 年に Karl Sims によって発表された[1, 2]。これらの研究では、Virtual Creature と呼ばれる人工生命が、3 次元物理シミュレーション空間において、進化的・自律的に身体と身体に合わせた行動の獲得を行った。Virtual Creature の動作は人工ニューラルネットワーク (Artificial Neural Network: ANN) によって制御されている。そのニューラルネットワークと身体の生成規則は、遺伝的アルゴリズム (Genetic Algorithm: GA) によって決定されている。参考文献[1]では特定のタスクを実現する行動の獲得に関する研究を行っている。タスクとしては歩行やジャンプの獲得などが採用されており、身体に合わせた歩行やジャンプなどが獲得された。参考文献[2]では、2 体の Virtual Creature による競争下でのシミュレーションが行われており、共進化による複雑な進化の様子が示された。

Karl Sims の研究は大変衝撃的であり、現在でも後続というべき研究が多数発表されている。例えば、Artificial Life 10 にて発表された論文では、3 件の論文が例として挙げられる。参考文献[11]では、フリーの物理シミュレーション用ライブラリである Open Dynamic Engine (ODE) [9, 10]を用いた Karl Sims の研究の再実験を行っている。また、参考文献[12]でも、ODE を用いて Karl Sims

の研究の再現を行い、投擲行動の進化的獲得について実験を行っている。参考文献[13]でも同様に、3次元物理環境を ODE で実現することで、Karl Sims の研究よりもより厳密な物理環境での実験を可能であると主張されている。また、続く参考文献[14, 15]では、その環境下で、NeuroEvolution of Augmenting Topology (NEAT)[16]と呼ばれる手法を拡張した手法を提案・適用し、自律的な行動獲得を実現している。

## **2.2 NeuroEvolution of Augmenting Topology (NEAT)**

NEAT[16]は 2002 年に Kenneth らによって発表された ANN の一種であり、誤差逆伝搬法と同等以上の問題解決能力を持つと主張され、現在多くの研究で利用されている。従来の ANN では、荷重を変更することで学習を行ってきたが、近年では荷重と ANN の構造を同時に変化させる手法が提案されている。NEAT もそのような手法の一種である。NEAT の特徴としては、以下の 3 点が挙げられる。

- 1) 異なる構造を持った ANN 間の交叉手法の採用
- 2) Speciation と呼ばれる構造の保護手法
- 3) 最小の構造 (2 層 ANN) から始めて、徐々に構造を大きくしていく点

図は NEAT の進化例を表している。最初は 2 層から始まり、荷重の値の変化、ノード間のリンクの追加、中間層へのノードの追加などを行い、徐々に構造を複雑なものへと変化させていく。



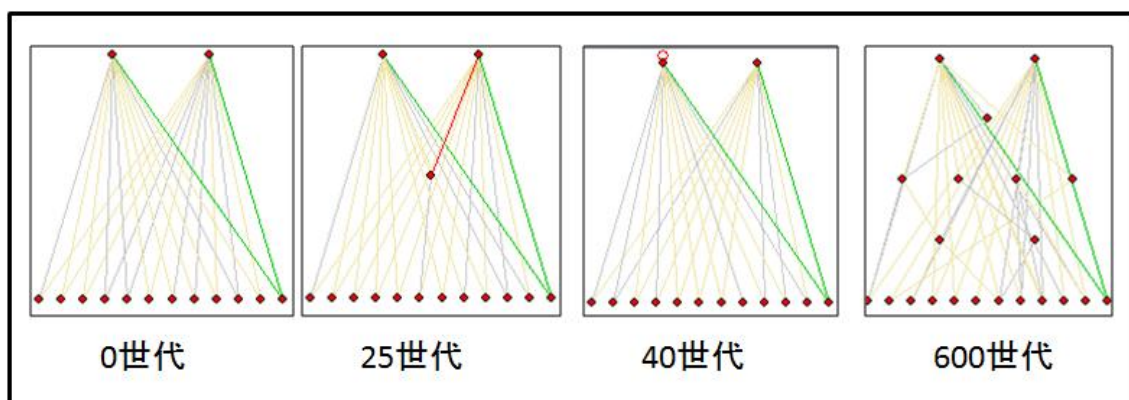


図 2.1 NEAT 手法における ANN の進化例

## 2.3 Evolutionary Robotics

Evolutionary Robotics[7]の分野では、実機ロボットの制御を人工ニューラルネットワークによって行い、そのニューラルネットワークに対して遺伝的アルゴリズム[6]を適用させることで進化させている。即ち、ロボットのセンサ情報をニューラルネットワークへの入力とし、ニューラルネットワークの出力値をロボットのモータの動作量とすることでロボットの制御を行い、制御の結果を評価し、最も良い評価を行った個体を使って次世代を生成する。ロボット単体の制御[17, 18, 19]、複数のロボットの共進化[20, 21, 22]、複数台の協調動作の獲得[23-28]まで幅広く適用が行われており、多大な成果が挙げられている。

## 2.4 Evolutionary Design

これらの研究では、構造物の設計図の進化的な獲得を行っている[3]。この構造物の素材としてはレゴブロックが想定されており、3次元シミュレーション環境において、レゴブロック同士の結合部に働く力が計算され、設計図の構造物が実現可能かどうか検証される。ブロック同士がどのように接続されるかという設計図は遺伝子として表現され、遺伝的アルゴリズムによって進化が行われ

る。図 2.4 は遺伝子の表現型の例を示している。評価関数は、目的とする構造物の特徴を評価する内容になっている。例えば、「できるだけ長さを持った構造」や「できるだけ重いものを保持できる構造」といった評価によって進化を行う事で、図 2.2、図 2.3 に示すような長い橋や重荷に耐えうる台などの構造が獲得されている。しかし、これらの研究では、あくまで構造物の設計の進化的な獲得を扱っており、実際にどうやってこれらの構造物を組み立てるかという設計手順については、獲得されていない。即ち、人間が考えて組み立てる必要があった。

そこで、参考文献[4, 5]では、Turtle と呼ばれるエージェントによる構造物構築行動の獲得を目的として研究が行われている。Turtle エージェントはプリンタヘッドのように空間上の X-Z 座標上の自由な位置に移動し、ブロックを設置することが可能である。図 2.5 に示すような行動を組み合わせる事で、図 2.6 に示すような構造物の構築手順の獲得を実現している。

しかし、これらの実験では、構築物の構築主体である Turtle エージェント自体の身体的特徴については考慮されていない。しかし、実際の生物と鑑みてみると、構造物と身体的特徴には必ず関係がある。環境、構造物、身体的な特徴、これらの関係を考慮する事でより高度な構造物の構築行動が獲得可能ではないかと考えられる。

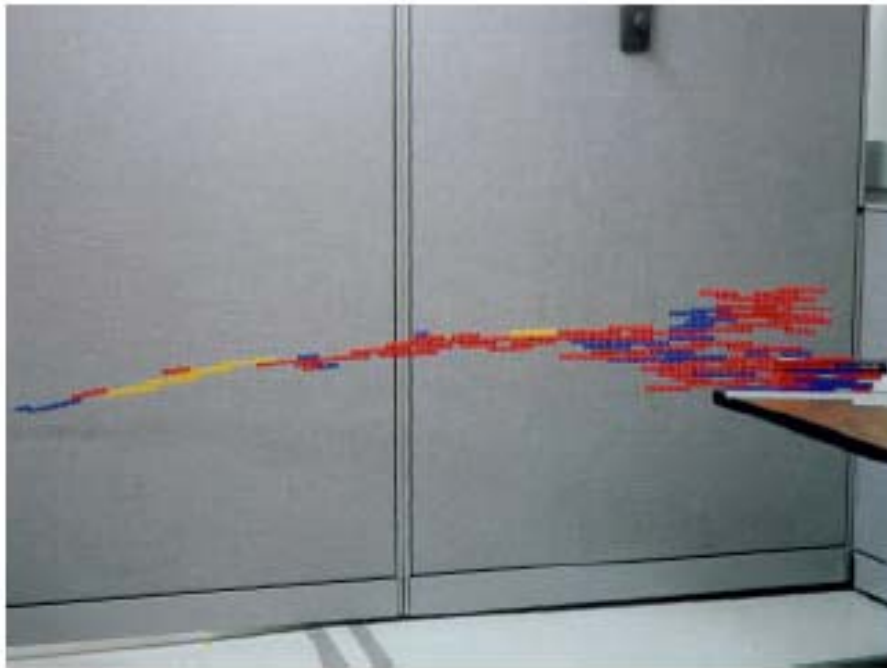


図 2.2 獲得された構造物の例 1



図 2.3 獲得された構造物の例 2

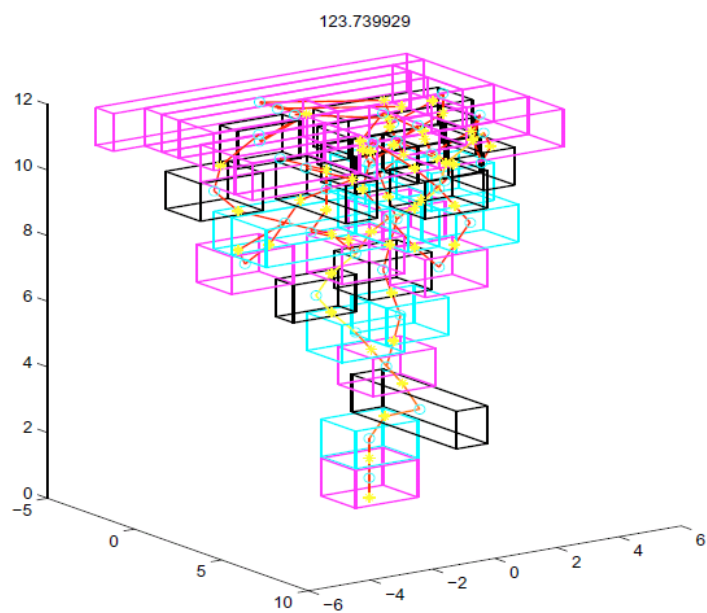


図 2.4 遺伝子から変換された構造物の設計図の例

※図 2.2～2.4 の出展は、参考文献[3]

Table 1. Parameterized Assembly Instructions

Instruction	Parameters
(M)ove	+2, +1, -1, -2
(R)otate	+90, -90, +180
(P)ut Brick	(a)head, to (r)ight, to (l)eft, (b)ehind
Put (S)caffolding	(a)head, to (r)ight, to (l)eft, (b)ehind
(T)ake Brick	(none)

図 2.5 Turtle エージェントの行動表

With Scaffolding					
Final					
Fitness	84%	80%	95%	90%	100% (hand-built)
Length	30	54	58	43	65

図 2.6 Turtle エージェントによって構築されたアーチの例

※図 2.5、2.6 の出展は、参考文献[4, 5]

## 2.5 2次元上でのクモの巣の構築行動に関する研究

構造物の構築を行う動物の構築行動獲得に関する研究では、主に数理生物学分野におけるアリやハチなどの社会性昆虫に関する研究が圧倒的に多い。その一方で、クモのような単体で構造物を構築する動物に関する研究はあまり多くはない。

参考文献[29]では、クモの造網行動をルールとして表現し、そのルールのパラメータを遺伝的アルゴリズムで最適化することで、進化的な造網行動の獲得を行っている。また、シミュレーションによって獲得された巣と、実際のクモが実験環境で張ったクモの巣との比較を行い、モデルの正当性について検討を行っている。ただし、彼らの研究では、クモ自体の身体的特徴や身体と環境との物理的な相互作用などについては、考慮されていない。

## 2.6 社会的ジレンマに対するゲーム理論とエージェントアプローチ

個人的合理性と社会的合理性が対立する状況における解決策の研究としては、従来、ゲーム理論的手法が適用されてきた。これは、1980年に Dawes[31]が囚人のジレンマ問題の一般化による定式化によって社会的ジレンマを表した事端起を発する。社会的ジレンマは、個人的合理性と社会的合理性が対立する状況下において、個人的合理性が優先された結果、社会的非合理が発生する状況を示す。社会的ジレンマ問題は、非協力ゲームの一種である繰り返しN人囚人のジレンマとして表現される事が一般的である。非協力ゲームとは、プレイヤーの間に行動選択に関して拘束的合意が成立しない事を前提として考えるゲーム状況である。即ち、プレイヤーは個人的合理性にのみ従って行動を決定し、社会的な合理性について考慮しない事を示す。これらの前提条件のもとで、社会

的ジレンマ問題は以下のように表現される。

N 人のプレイヤーが協調行動 C か非協調行動 D のどちらかを自身の合理性に基づいて選択を行う。社会的ジレンマ問題の場合は、非協調行動が支配戦略であり、全てのプレイヤーが非協調行動を選択する状況が均衡状態となる。しかし、その状態で得られる利得は、全てのプレイヤーが協調行動を選択したときよりも常に小さくなる。

上述のように、社会的ジレンマ問題に対しては、ゲーム理論によるアプローチが従来的であったが、近年、エージェントベースアプローチによる研究が盛んになっている。エージェントベースアプローチは、状況に応じて意思決定を行う限定合理的な主体を設計することで、ボトムアップ的にシステムを構築する手法である。この手法が採用される理由として、(1)ゲーム理論におけるプレイヤーとエージェントの親和性が高い事、(2)数理モデルによる解析的な分析が難しい複雑な設定においても適用可能である事、(3)動的な環境においても適用が容易である事、といった3点が考えられる。

社会的ジレンマ問題へのアプローチは、構造改革型アプローチと態度変容型アプローチの2種類に大別できる。

構造改革型アプローチは、プレイヤーの非協力的な行動に対して外部から圧力を加える事で問題の社会構造を変更するアプローチである。具体例としては、プレイヤーの活動に対して課税や罰金などを設定し、プレイヤーが本来得られるはずの利得に変更を加える方法が挙げられる。構造改革型アプローチの利点は、適切な構造改革を行う事ができれば、全てのプレイヤーに協力行動を促す事が可能な点である。欠点としては、その適切な構造改革を見つける事が難し

い点である。

態度変容型のアプローチは、プレイヤー同士の連帯感や信頼感・モラルの構築によって、プレイヤーの合理性の種類を変えるアプローチである。社会的ジレンマ問題におけるプレイヤーの個人的合理性とは、プレイヤー自身の利得を最大化する事を目的する行動方針であると言える。態度変容型アプローチの具体例としては、プレイヤーの合理性を全体的な合理性を考慮するようなものに変更を加えるアプローチが挙げられる。このアプローチの利点としては、いったん集団への連帯感が構築されると協調が持続する事である。欠点としては、プレイヤーの意思決定に対して強制力を持たない為、非協力者がなくなることである。

構造改革アプローチの例としては、山下らによって提案されたメタエージェントによる課税戦略[33]が挙げられる。山下らは *Iterated Multiple Lake Problem* に対して適用を行い、先行研究では、共有地の悲劇への適用を試みている[13]。

## 2.7 生物学での関連研究

参考文献[8]では、クモの生態について詳しく記載されている。クモの身体的構造や特徴から、行動の種類、その特徴、それらに関する実験など、その範囲は多岐に渡っている。本研究に関連する部分だけを抜粋した、クモの身体的特徴、行動的特徴は第5章にて、その詳細を述べる。

### 3. メタ行動の導入の枠組み

本章では、メタ行動導入の枠組みについて定義を行い、そのモデル化の議論を行う。直接的な行動は、報酬関数によって直接的に学習・進化するが、それに対して、直接行動のサポートを行う間接的行動であるメタ行動では、ある種のコストを通じて間接的に報酬関数を変更し、報酬を得ると考えられる。その詳細について論じ、モデル化を試みる。また、本研究で扱うメタ行動導入の題材である、共有地の悲劇と造作物構築行動についてメタ行動導入のモデルを用いて表現を行い、各題材の特徴と手法の対象領域について明確化する。

#### 3.1 メタ行動の導入の枠組みのモデル

本研究では、行動の種類について2種類の行動に焦点を当てている。1つは、直接的な行動、もうひとつは、直接的な行動をサポートするような間接的な行動である。本研究では、前者を直接的行動、後者をメタ行動と呼ぶ事としている。この名称には多分に誤解を生じる可能性があるとは考えられるが、間接的行動の中でも直接的行動をサポートする間接的行動である点、ある行動の為の行動であるということから行動間の階層性が考慮できるという点から、この名称を採用している。メタ行動は、環境に対してなんらかの変化を及ぼすことで直接的行動の効率性や効果を上昇させる間接的な行動を指す。

図 3.1 にメタ行動を含む環境について示す。直接行動のみを考慮した環境の場合、エージェント **Agent** は環境からのセンシング **S** を元に、行動 **A** を決定し、実行する事で、自身を含む環境 **Env** から報酬 **R** を獲得する事が出来る。即ち、この環境下で獲得される報酬は以下の式で表す事が出来る。



$$R = Agent (Env, A) \quad (3.1)$$

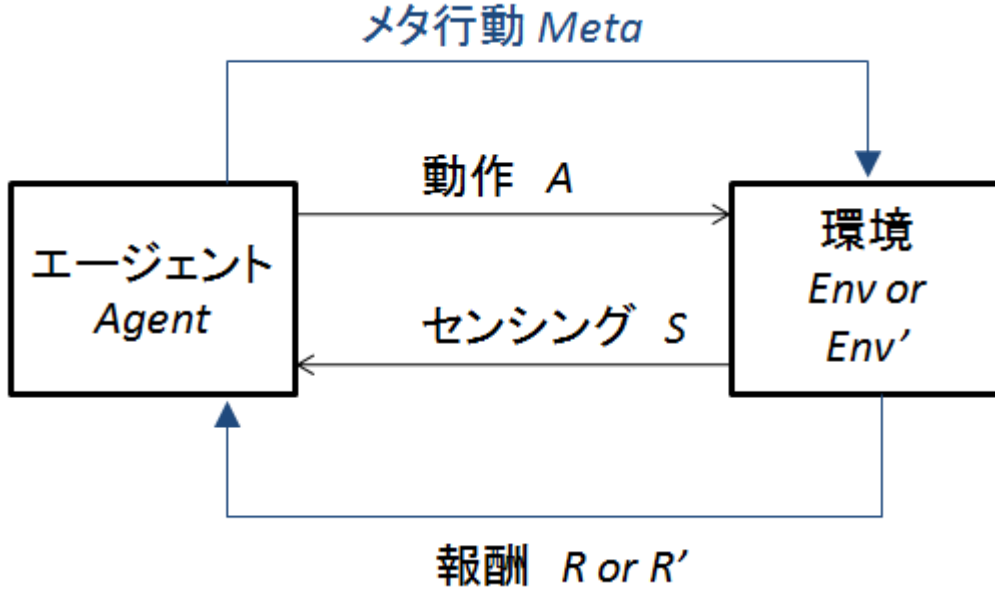


図 3.1 メタ行動導入前後の環境について

ここで、環境の特性を変更しうるメタ行動を考慮した場合を考えてみる。メタ行動が実行された場合、エージェントから環境に対する能動的な変化が与えられ、環境が  $Env$  から  $Env'$  の状態へと変化する。変化した結果として獲得される報酬  $R'$  は以下の式で表す事が出来る。

$$R' = Agent(Env', A, Meta)、但し \quad Env' = Agent(Env, Meta) \quad (3.2)$$

即ち、

$$R' = Agent(Agent(Env, Meta), A, Meta) \quad (3.3)$$

である。

エージェントは常に利得を最大化させる利己的なエージェントであると仮定すると、エージェントは動作 A かメタ行動 Meta のいずれかから自己の利益を最大化することが出来る行動を選択すると考えられる。

実際の動物による捕獲行動を例にとって考えてみると、直接的行動は、追いかけて捕まえるような直接的な捕獲行動、そのメタ行動は造網行動のような間接的な捕獲行動であると考えられる。直接捕獲をするためには、獲物よりも身体的特徴に優れている必要がある。一般的に、捕獲者の移動速度や体格などが獲物よりも大きくなければ獲物を捕獲することはできず、逃げられてしまう。一方、造網行動によって巣を構築することは直接捕獲に比べてより多くのコストを必要とするが、直接捕獲では捕まえないような獲物を捕まえる事も出来る。

コストに関して考えてみると、直接捕獲に比べて、間接捕獲は直接捕獲をサポートする為の行動や環境を変更する為のコストが必要になってくる。即ち、環境を変更する間は、獲物を捕まえる事が出来ない事を考えると、効率性の低下が発生すると考えられる。直接捕獲のコストを  $C_D$ 、間接捕獲時のコストを  $C_I$  とすると、直接捕獲と間接捕獲における効率は以下のように考えられる。

$$\text{直接捕獲の効率} \quad E_d = R / C_D \quad (3.4)$$

$$\text{間接捕獲の効率} \quad E_i = R' / C_I \quad (3.5)$$

この時、直接捕獲の効率が低い状況、即ち  $E_d > E_i$  であれば直接捕獲が選択され、 $E_d < E_i$  であれば間接捕獲が選択されると考えられる。

間接捕獲は初期の段階ではコスト・時間ともに直接捕獲よりも多く必要となる場合が多い為、即ち、 $C_D \leq C_I$  であると考えられる為、安定的・継続的に選択

されにくいと考えられる。また、間接捕獲の効率が直接捕獲よりも高く、最終的に得られる利得が大きくなければ間接捕獲を進化的に獲得することは難しいと考えられる。

### 3.2 メタ行動の導入の枠組みを用いた各題材の表現

本研究では、2種類の題材に対してメタ行動の導入を行う。1つ目の題材は、ゲーム理論の問題のひとつで、有限な共有資源のシェアリングに関するジレンマを取り扱った問題である共有地の悲劇である。この問題では、複数のエージェントが自身の利得を最大化させる為に、共有資源の使用量である活動度の決定という行動を行う。しかし、この問題では、共有資源の最大量は決まっており、全体の活動度が高い程、共有資源から得られる利得が低下していくという特徴的な環境設定となっている。その為、全員が利己的に最大活動度を選択し続けると最低限の利得しか獲得できないが、利己的なエージェントは活動度を下げる事ができない、というジレンマ的状况である。本研究では、共有地の悲劇に対して、環境を表す利得関数に対して変更を加える行動である課税行動というものをエージェントに対して導入している。即ち、直接行動である活動度の選択か、メタ行動である課税行動のどちらか、自身の利得が最大化する行動を選択するようにエージェントに変更を加えている。

一方、もう一つの題材である捕獲用の造作物構築行動の獲得では、捕獲者エージェントによる獲物エージェントの捕獲行動の獲得を題材としている。捕獲者エージェントは、直進、右旋回、左旋回、造作物の材料であるブロックの設置といった基本的な行動を用いて、直接行動である追跡による直接的捕獲か、メタ行動である造作物を用いた間接的捕獲の獲得を行う。

共有地の悲劇は、確定的なルールによってゲーム全体の挙動が予測可能であるという特徴がある。また、エージェントの行動に対してコストが考慮されず、実行によって即座に利得が獲得されるという環境であり、行動のプロセスではなく、動作が評価対象となっている。エージェントの行動は、直接行動かメタ行動かのいずれかを選択するという形式になっている。それに対して、捕獲用の造作物構築行動の獲得では、確定的なルールによってシミュレーションの挙動が決定されておらず、ランダム要素を含む為、環境の挙動の予測は難しい。また、複数の動作を組み合わせて直接捕獲や間接捕獲の一連のプロセスを獲得する必要があり、獲得したプロセスが造作物の評価を通じて間接的に評価されていると言える。行動についても、直接行動かメタ行動の選択ではなく、基本的な行動の組み合わせを用いて直接行動やメタ行動を獲得する。以上の特徴をまとめたものを表 3.1 に示す。以上の様に、本研究では、メタ行動の導入の枠組みを適用する問題とその対象領域を段階的に広げていく事で、研究を進めていく。

表 3.1 本研究で対象とする題材の特徴

	共有地の悲劇	造作物構築行動の獲得
環境の性質	確定的 予測可能	不確定的 予測困難
評価の対象	行動単体	行動のプロセス
行動の獲得方法	直接行動かメタ行動かの 選択	基本的動作からの獲得

## 4. ゲームエージェントにおけるメタ行動

本章では、ゲーム理論におけるメタ行動について議論を行い、メタ行動による枠組みの理論的な問題における有効性について検証を行う。本論では、ゲーム理論の題材の中でも、メタ行動との親和性が高いと考えられる共有地の悲劇について説明を行い、共有地の悲劇にたいしてメタ行動を適用することで問題解決を図る手法について議論を行う。

### 4.1 共有地の悲劇

#### 4.1.1 概要

共有地の悲劇[30]は、1968年に生物学者 Garret Hardin によって発表された寓話で、環境問題への警鐘であった。寓話の内容としては以下の通りである。

共有の牧草地に対して複数の牛飼いが好きなだけ牛を放牧する事で過放牧が発生する。しかし、常に自身の利得を最大化させようとする牛飼い達は、自分ひとりだけが放牧している牛の数を減らす事が出来ない。そうして、過放牧の状態が続いた結果、牧草地の荒廃が進み、全ての牛飼いは利益を得る事が出来なくなった。

この寓話の特徴は、牛飼い達の利益という個人的合理性と共有の放牧地の保護という社会的合理性が対立する状況において、個人的合理性を優先した結果として、社会的非合理が発生するという点である。現代社会における環境問題や資源枯渇問題などの多くは、この社会的ジレンマ構造を持っているといわれている。具体的な例としては、二酸化炭素による地球温暖化問題や、フロンガスによるオゾン層破壊の問題、化石燃料の枯渇問題などが挙げられる。

以上が、共有地の悲劇の設定である。共有地の悲劇では、このままの設定では悲劇的状況を回避する事が出来ないことが多くの研究によって示されている。そこで、共有地の悲劇に対してどのような拡張を行えば、悲劇的状況を回避可能であるかという研究が行われてきた。

#### 4.1.2 共有地の悲劇の定式化

共有地の悲劇に対する定式化は、繰り返し  $N$  人囚人のジレンマ問題による一般化[35]に端を発する。その後、社会的ジレンマ問題のモデルについても様々な研究が行われてきた。宮西ら[32, 36]は、[35]に示された利得関数を用いてモデル化を行っている。他のモデルとの違いとしては、(1)  $N$  人囚人のジレンマ問題としてではない定式化が行われている点、(2) 非協調行動に非協調の度合いが設定されており、エージェントの選択が 2 択ではない点、以上の 2 点が挙げられる。以下は、そのモデルを用いた共有地の悲劇の定義である。

このゲーム問題では、 $N$  体のエージェントはそれぞれ、自身の利得を最大化する事を目的として活動度の選択を行う。活動度は共有資源の消費度合いを表し、数値が高い程得られる利得が多い。活動度が 0 の場合は協調行動、それ以外の場合は裏切り行動を選択したと考えられる。行動の選択によって得られる利得は自身の行動のみならず全エージェントの選択に依存している。共有地の悲劇は非協力ゲームの一種であり、自身が行動を決定する際に他のエージェントの行動を知る事ができない。即ち、エージェントは  $TA$  の値を見積もった上で自身の行動を決定する必要がある。しかし、このゲームにおいては、どのような  $TA$  の値においても、高い活動度ほど多くの利得を得られる。従って、全ての

エージェントが個人的合理性に従う場合、推定される TA の値は常に最大値であり、その場合においても自身も最大活動度を選択せざるを得ない。

即ち、エージェントの数を  $N$  とした場合、上記の設定は以下のように表される。

活動度  $Activity = \{act_j \mid 0 \leq j \leq M\}$

エージェント  $i$  の活動度  $a^i \in Activity$

エージェント  $i$  の得られる利得

$$Payoff(a^i, TA) = a^i(M \times N - TA) - 2a^i \quad (4.1)$$

但し、 $TA = \sum_{i=1}^N a^i$  は全エージェントの活動度の合計を示している。

以下に、 $N=4$ 、 $M=4$ 、 $Activity=\{0, 1, 2, 3\}$ である場合の利得関数の例を表 4.1 に示す。

$$Payoff(a^i, TA) = a^i(16 - TA) - 2a^i \quad (4.2)$$

表 4.1 共有地の悲劇における利得関数の一例

		Total Activity (TA)												
		0	1	2	3	4	5	6	7	8	9	10	11	12
$a^i$	0	0	0	0	0	0	0	0	0	0	0	-	-	-
	1	-	13	12	11	10	9	8	7	6	5	4	-	-
	2	-	-	24	22	20	18	16	14	12	10	8	6	-
	3	-	-	-	33	30	27	24	21	18	15	12	9	6

また、表をグラフ化したものを図 4.2 に示す。

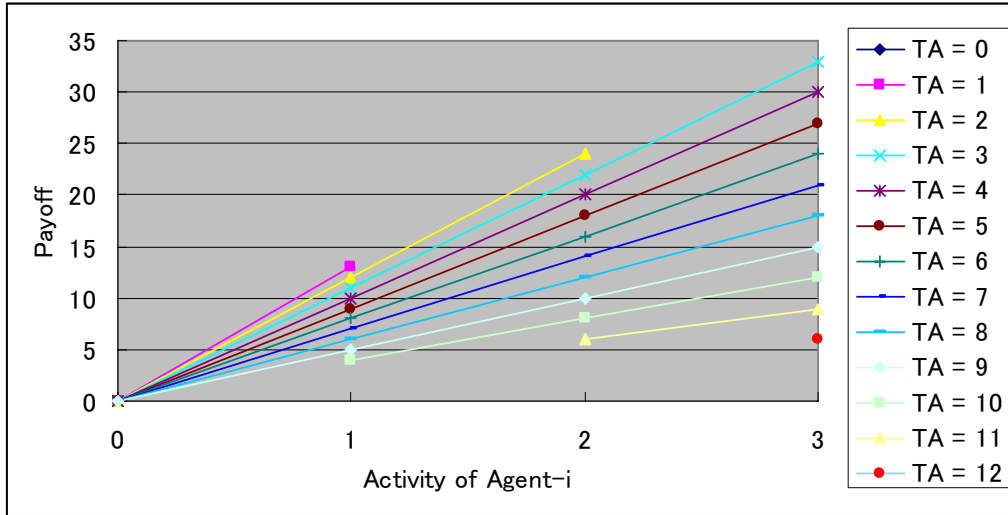


図 4.1 共有地の悲劇における利得関数の一例

以上の例からわかる通り、どの状況においても活動度が高いほど得られる利得は高く、TA の値が低いほど得られる利得は高い。即ち、この利得関数の特徴は以下のように現される。

$$\begin{aligned}
 \text{Payoff}(\text{act}_{j+1}, \text{TA}) &> \text{Payoff}(\text{act}_j, \text{TA}) \\
 \text{Payoff}(a^i, \text{TA}) &> \text{Payoff}(a^i, \text{TA}+1)
 \end{aligned}
 \tag{4.3}$$

社会的ジレンマ問題においてジレンマ状況を回避する為の方法として提案されている手法は大別すると以下の 2 つである[34]。一つは、個人的合理性を他の種類の合理性に変更する手法、もう一つは利得関数を変更する事で問題構造を変更する手法である。前者は、エージェントの意思決定方法に対して、他者との関係や社会的な利益を評価するような変更を加える手法であり、エージェント間の協調行動の誘発が期待できる。しかし、行動の選択に対する外部的な強



制力が無いために、協調行動グループを犠牲にして自身の利益を獲得するフリーライダーの出現を抑える事が出来ない。後者の手法は、報酬や罰金などによって利得構造の変更を行う手法である。エージェントは、変更後の利得構造を基に個人的合理性を満たすような行動を選択する。即ち、悲劇的状况を回避可能な報酬制度や罰金制度を作り込むことができれば、個人的合理性を変更することなく悲劇的状况を回避する事が可能である。また、非協調行動に対する外部的な強制力によってフリーライダーを抑制する事も可能となる。しかし、この手法の問題点として、最適な報酬制度や罰金制度、それらの設定をどの様に発見するかが問題となってくる。

山下らの研究では、利得構造の変更によってジレンマ状況を回避する手法の一種である「メタエージェントによる課税戦略」の提案と導入を行っている[33]。また、本研究とその先行研究では、メタエージェントによる課税戦略に対して、メタエージェント化機能の導入を行っている[32, 36]。

第 3 章で示したメタ行動導入後の枠組みで考えた場合、メタ行動導入は後者の手法、問題構造を変更する手法の一種であると言える。拡張を加えた共有地の悲劇におけるエージェントは、通常の行動に加えて環境に値する利得構造の変更を行う「報酬や罰金を用いる」というメタ行動を選択可能であり、それをうまく用いる事で、状況を以前よりも改善することが可能である。山下らの研究では、メタ行動を選択したエージェントはメタエージェントと呼ばれ、自身の持つ課税戦略を元に利得構造を変化させる事ができる。以下では、その詳細について述べる。

#### 4.1.3 メタエージェントによる課税戦略

この戦略では、社会的ジレンマ問題の利得構造に変更を加えるために、エージェントの消費行動に対して課税を行う。即ち、各活動度に対して課税を行い、エージェントが本来得られる利得を減額させる。メタエージェントによって行われる各活動度に対する課税は、各エージェントの持つ課税プランとして表現される。メタエージェントの利得は、課税プランを実施して得られる税金によるものなので、各活動度を選択したエージェントの数  $N(a_j)$  と  $a_j$  に対応する課税値の積の合計で表される。また、メタエージェントも個人的合理性に従い税金を最大化するように課税の設定を行う。但し、メタエージェントの課税値の上限は、本来の利得よりも低い値であるとする。

メタエージェント  $k$  の課税プラン

$$LP^k = \{lv_j^k | 0 \leq j \leq M\} \quad (4.4)$$

*where*  $0 \leq lv_j^k \leq \text{Payoff}(a_j, TA)$

メタエージェント  $k$  の利得

$$\text{Revenue}^k = \sum_{j=0}^M N(a_j) * lv_j^k \quad (4.5)$$

エージェント  $i$  の利得

$$\text{Reward}^i = \text{Payoff}(a^i, TA) - lv_{a^i}^k \quad (4.6)$$

メタエージェントが1体のみの場合、個人的合理性に従うメタエージェントは出来る限り税金を高める課税プランを設定しようとする。そのような事態を回避する為に、複数のメタエージェントによる競争の導入を行っている。即ち、

各メタエージェントの課税プランから最小課税値で構成される社会的課税プランを作成し、最小課税値を提案したメタエージェントのみが税金を得られるように設定した。これにより、メタエージェント間の競争が発生し、課税値の利己的な上昇を抑制する事が可能となった。社会的課税プランの導入に関する変更を加えた結果を以下の式として示す。

社会的課税プラン

$$SLP = \{lv_j^{\min} | 0 \leq j \leq M\} \quad (4.7)$$

where  $lv_j^{\min} = \min(lv_j^1, \dots, lv_j^N)$

メタエージェントの利得

$$Revenue^k = \begin{cases} \sum_{j=0}^M N(a_j) * lv_j^{\min} & (if \quad lv_j^{\min} = lv_j^k) \\ 0 & (if \quad lv_j^{\min} \neq lv_j^k) \end{cases} \quad (4.8)$$

エージェントの利得

$$Reward^i = Payoff(a^i, TA) - lv_{a^i}^{\min} \quad (4.9)$$

以下の図は、課税プランとエージェントの得られる利得の関係を表したものである。

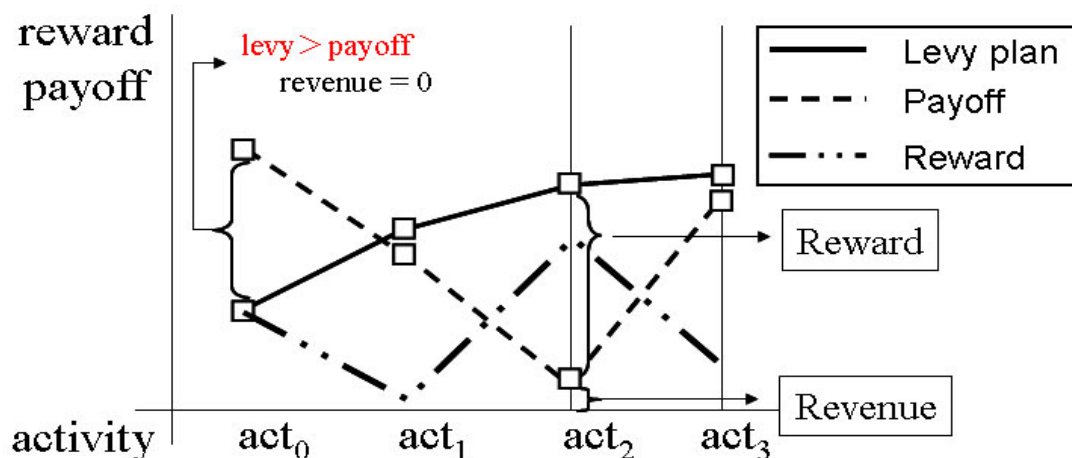


図 4.2 課税プランとエージェントの利得 (Reward) の関係

グラフにおける横軸はエージェントの活動度を示し、縦軸は利得と課税値を示す。グラフに示すような課税プランが提示されている場合、エージェントが得られる利得 **Reward** は上記のように表す事ができる。活動度  $act_0$  においては、課税値が利得よりも大きいので、メタエージェントは税金を取り立てる事が出来ない。他の活動度では、**Payoff** から課税値を差し引いた値が **Reward** になっている。エージェントはこの **Reward** をもとに、個人的合理性を満たす活動度を選択する。この場合、活動度  $act_2$  を選択した時に得られる利得が最も高いので、エージェントは活動度  $act_2$  を選択する。

以上の仕組みを用いる事で、メタエージェントの課税プランによっては悲劇的状况を回避することが可能ではないかと考えられる。メタエージェントによる課税戦略を採用する際に考慮しなければならない点として、以下の2点が考えられる。

1. 適切な課税プランをどのように設定すべきか。
2. メタエージェントを何体導入すべきか

適切な課税プランを設定する事ができれば、社会的ジレンマ状況を回避できる可能性が出てくる。問題を分析する事によって適切な課税プランを埋め込む事は可能である。しかし、エージェントベースシミュレーションの目的の一つは、自律的に問題を解決する事である。従って、適切な課税プランを自律的に獲得できる事が望ましい。その為に、進化的手法の導入を行い、課税プランに対して遺伝的アルゴリズムの適用を行っている。メタエージェントは複数ある遺伝子のなかから最も評価の高い遺伝子を課税プランとして採用する。遺伝的アルゴリズムにおける評価関数の詳細については後述する。

メタエージェントを何体導入すべきかについては、問題の設定によって変わる事が予想される。この問題についても自律的に解決される事が望ましい。本研究では、エージェントから状況に応じてメタエージェントが選出される事で、この問題を解決している。次節では、その方法について詳細を述べる。

## **4.2 メタエージェント化機能の導入**

メタエージェントをどのタイミングで、何体導入すればよいのかという問題に対して、本研究はメタエージェント化機能を導入する事で解決を行った。エージェントが選択する活動度の一つとして、メタエージェントとして振舞う事を追加している。即ち、エージェント自体を拡張している。拡張の際、共有資源を利用する活動度を選んだエージェントをプレイヤー、課税を行うエージェントをメタエージェントと定義した。従って、エージェントの行動は以下の様に定義される。

$$\text{エージェントの行動 } a^i \in \{A_{player} \cup A_{meta}\} \quad (4.10)$$

$$\text{プレイヤーとしての活動度 } A_{player} = \{a_j | 0 \leq j \leq M\} \quad (4.11)$$

$$\text{メタエージェントとしての活動度 } A_{meta} = \{a_{meta}\} \quad (4.12)$$

エージェントは、自身の行動を選択する際にプレイヤーとしての期待利得とメタエージェントとしての期待利得を比較して最も利得の高い活動度を選択する。期待利得は、上述の **Reward**、**Revenue** と同じ式を用いて計算される。

### 4.3 遺伝的アルゴリズムにおける評価関数について

先行研究では、どのような評価関数を設定すれば、適切な課税プランを獲得する事ができるかを調べるために、以下の評価関数の導入を行っている。

$$E_i = \left( \frac{\alpha \cdot Revenue_i + \beta \cdot Reward_i}{Reward_{i_{worst}}} \right)^a + \left( \frac{\sum_{j=1}^N Reward_j}{\sum_{j=1}^N Reward_{j_{worst}}} \right)^b \quad (4.13)$$

この評価関数は、パラメータ設定によってエージェントの個人的合理性と社会的合理性の評価の比率、エージェントの役割に対する評価の比率を変更する事ができる。第一項は、エージェントの個人的合理性に対する評価を表している。即ち、プレイヤーとしての利得とメタエージェントとしての利得の合計を用いた評価を行っている。第二項は、システム全体の社会的合理性に対する評価を行っている。このシステムにとって良い状態とは、ジレンマ状況を回避して全てのエージェントの利得を増加させる事である。従って、全てのエージェントのプレイヤーとしての利得の合計値を用いて評価を行っている。第一項、

第二項の分母は、エージェントの得られる最低の利得である。即ち、悲劇的状況が発生した場合にエージェントが得られる利得を示している。

この式のパラメータは  $\alpha$ 、 $\beta$ 、 $a$ 、 $b$  の4個である。 $\alpha$  はメタエージェントとしての利得に対する係数、 $\beta$  はプレイヤーとしての利得に対する係数である。パラメータ  $a$  は個人的合理性に対する係数、パラメータ  $b$  は社会的合理性に対する係数である。これらの値を調節する事で、様々な評価関数を用いる事が可能となる。例えば、 $a \gg b$  という設定にする事で、社会的合理性に関する項目を評価しない評価関数として扱う事が可能となる。

#### 4.4 拡張版の共有地の悲劇のシミュレーションについて

図 4.3 はシミュレーションの流れの概要を示したものである。また、PADで表現したものが、図 4.4 になる。

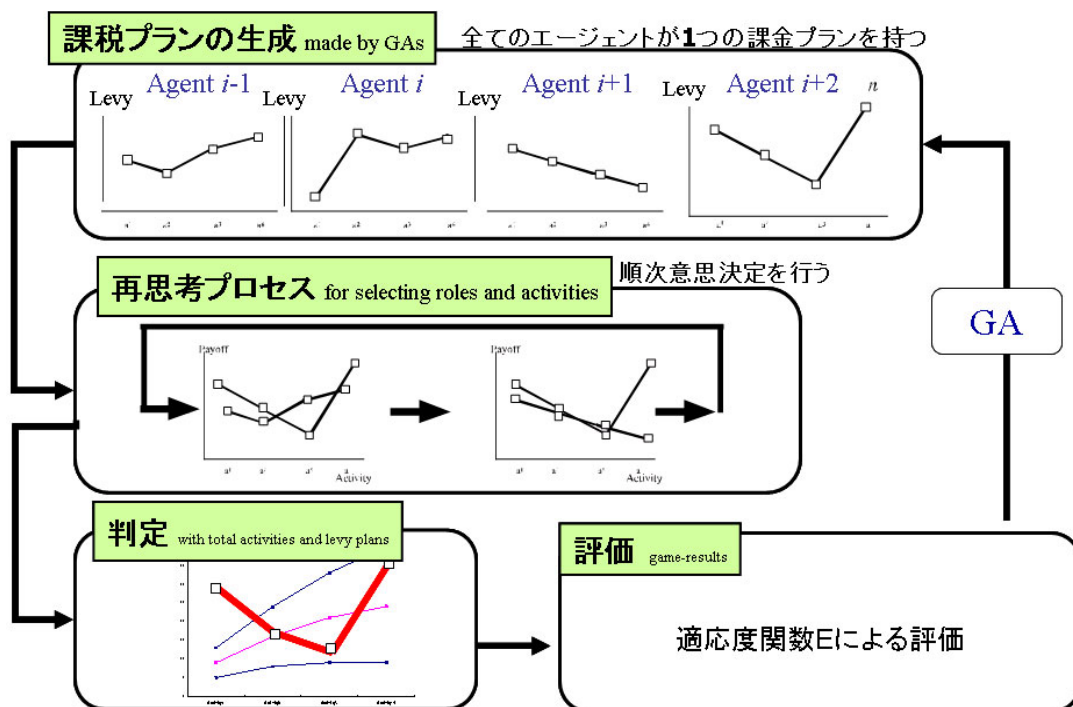


図 4.3 拡張版共有地の悲劇のシミュレーション概要図

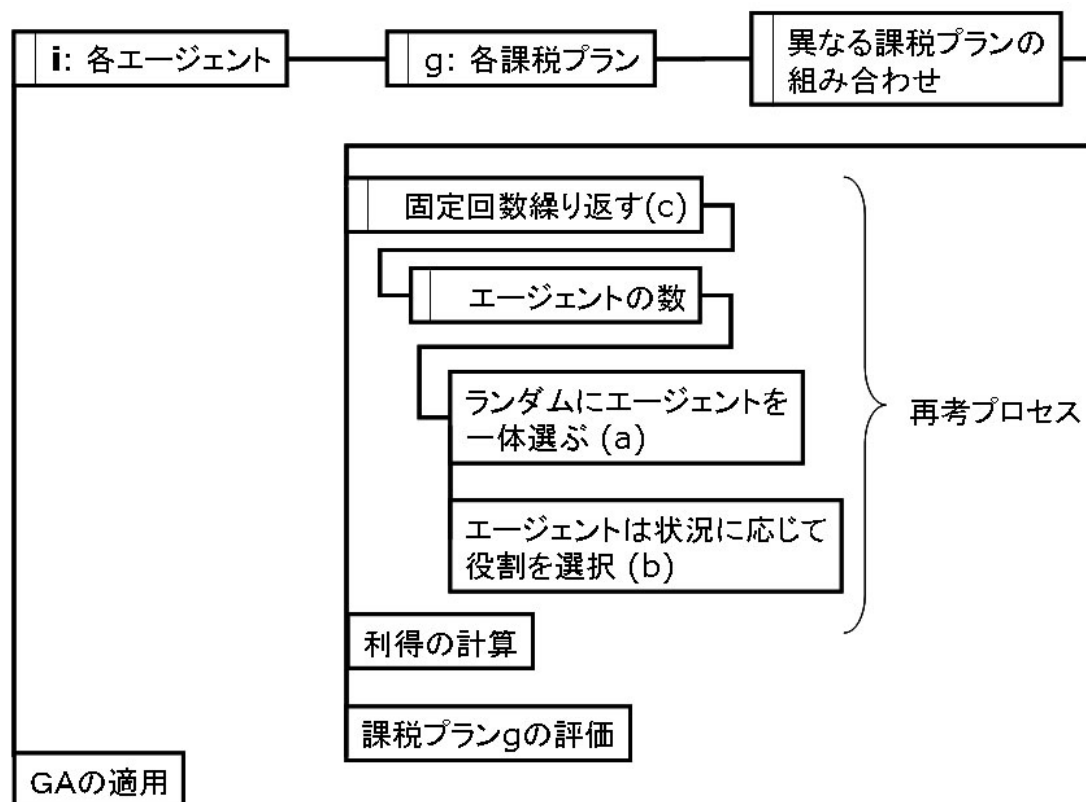


図 4.4 拡張版共有地の悲劇のPAD図

シミュレーションでは、まずエージェントの設定を行うことから始まる。ここで、課税プランの設定が行われるが、初期ステップの場合はランダムに課税プランを用いる。次のステップからは遺伝的アルゴリズムの遺伝的操作によって生成された課税プランを用いる。次の再考プロセスでは、エージェントの意思決定を行っている。このステップでは、ランダムな順番でエージェントに意思決定をさせるプロセスを固定回数繰り返す。これは、エージェント間の情報量の格差を埋めるためである。なぜならば、最初に選ばれたエージェントは他の全てのエージェントの活動度を推測して自身の活動度を選択しているのに対して、最後に選ばれたエージェントは推測なしで自身の活動度を選択できる。この情報的な格差を埋める為に、間隔をあけて意思決定を複数回繰り返してい



る。全てのエージェントの意思決定が終了した後に、エージェントが得られる利得の計算を行い、得られた利得を基に課税プランの評価を行う。最後に次のステップで用いる課税プランを遺伝的操作によって作成する。ここまでのシミュレーションのおおまか流れである。以下に課税プランの評価についての詳細を記載する。

再考プロセスからこの評価プロセスまでの一連の流れを通して、各エージェントの持つ一つの課税プランについて評価を行う事が出来る。しかし、遺伝的アルゴリズムを適用するためには、他の全ての課税プランについても評価を行う必要がある。そこで、全ての課税プランを評価するために、再考プロセスから評価プロセスまでの流れを遺伝子の数だけ繰り返す事になる。例えば、エージェントが10体、各エージェントが遺伝子を50個持っていたとすると、プロセスを $50 \times 10$ 回繰り返す事になる。因みに、課税プランの評価は他のエージェントがどのような課税プランを提案していたかによって変化する。そのため、このシミュレーションでは、ある課税プランの評価を行うために、他のエージェントが異なる課税プランを取った場合の評価も行い、平均をとっている。例えば、1つの課税プランにつき6回評価を行った平均値をその課税プランの評価値とする場合、以上の流れは、 $6 \times 50 \times 10$ 回繰り返される事となる。

#### 4.5 メタ行動導入の枠組みでの表現

図 4.5 は、これまでの共有地の悲劇の拡張に関する説明を、メタ行動の導入の枠組みを用いた共有地の悲劇のモデルとしてまとめたものである。今回の実験では、全エージェントが利得を向上させる事が出来る課税プランの自律的な獲得を行う。即ち、メタ行動の自律的な獲得に関する実験を行う。

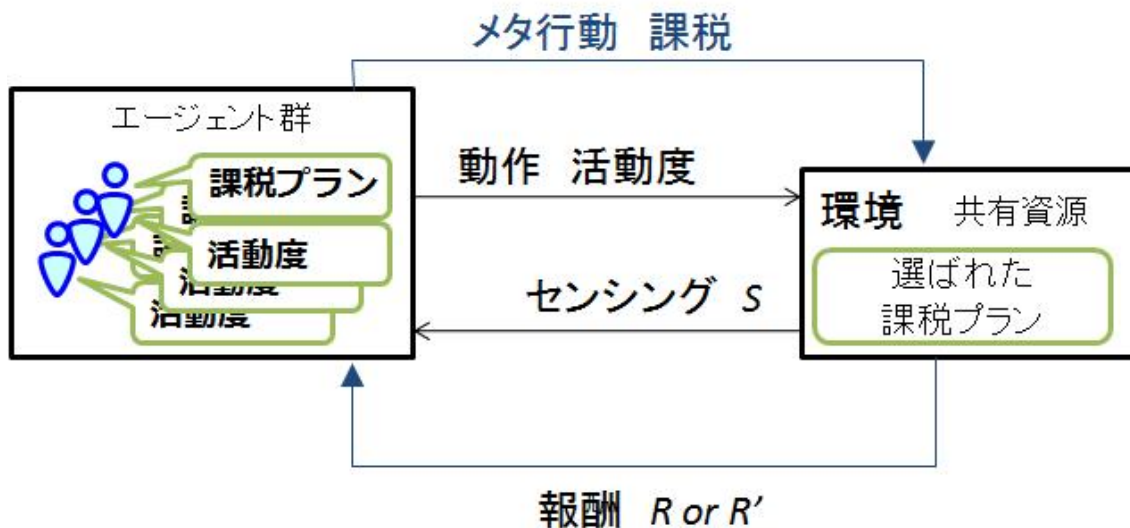


図 4.5 拡張版共有地の悲劇のメタ行動の導入の枠組みによる表現

#### 4.6 シミュレーション実験とその結果について

以下の実験パラメータを用いて、シミュレーションを行い、メタエージェントとプレイヤーの利得の変化について調べた。実験に関するパラメータについては、以下の通りである。

エージェントの数	12 体
活動度の最大レベル	6
初期活動度	プレイヤーとしての最小活動度 $a^i = a_0$
評価関数の定数	$\alpha = 0$
	$\beta = 2$
	$a = 0$
	$b = 2$

#### 遺伝的アルゴリズム関連

遺伝子の数	30
突然変異率	0.05%

再考プロセスの繰り返し回数 12 回

シミュレーションの結果を図 4.6 に示す。図は各世代における平均利得の遷移を示すものである。5 世代毎に平均を取ったプレイヤー全体の利得の平均（図中菱型）、メタ全体の利得の平均（図中四角）、プレイヤーとメタの利得を合わせた状態での平均利得（図中丸）、提案手法導入前の悲劇的状況におけるエージェント全体の利得の平均（図中三角）を示している。

悲劇的状況における利得は、式 4.1 の利得関数に実験パラメータを代入して求められた数値である。常に全エージェントが最低利得である 50 を獲得し続ける状況になっている。それに対して提案手法導入後のプレイヤーの平均利得は、常に最低利得である 50 を上回り、初期世代では 10 程度、最終的には 20 程度、平均利得を上昇できている。メタの平均利得を見てみると、変動は大きいもの

の、プレイヤーの利得を高い状態に維持しながら、自身の利得も高い状態で維持しており、適切な課税プランの獲得が出来ている事がわかる。メタエージェント 1 体とプレイヤー11 体の合計の平均利得を見ても、悲劇的状况を脱している事がわかる。

以上のように、課税行動というメタ行動の導入を行い、その適切な課税行動を獲得する事で、全体としての利得を上昇させ、悲劇的な状況を回避する事が可能であることを示す事ができた。

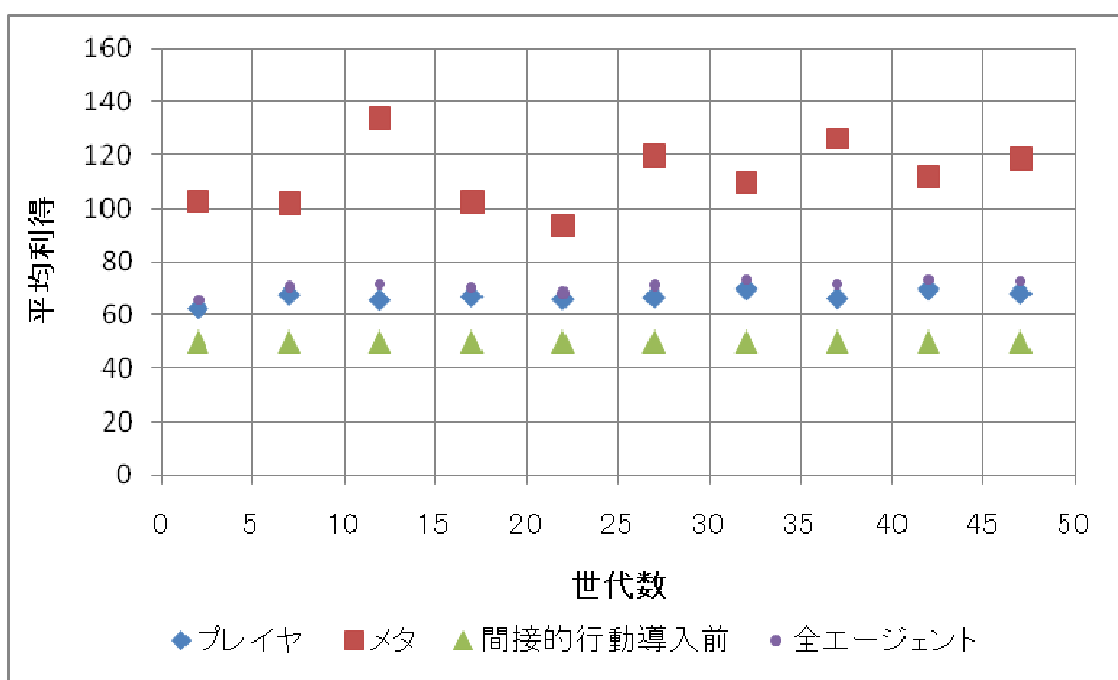


図 4.6 世代による平均利得の遷移

#### 4.7 拡張版共有地の悲劇におけるメタ行動の獲得について

ここで、このゲームにおける直接的行動とメタ行動とのコストについて比較を行い、第3章で提案したモデルの有効性について検証を行う。

このゲームの場合、プレイヤーになるエージェントの行動回数とメタになるエージェントの行動回数は一致するので、コストの比較は獲得された平均利得の比較と等しいことになる。図4.6を見てみると、直接的行動を行ったプレイヤーの平均利得は70前後であるが、間接的行動を行ったメタの平均利得は120前後になっている。即ち、適切なメタ行動が獲得された状況において、 $E_d \doteq 70 < E_i \doteq 120$  が成り立っている事がわかる。

#### 4.8 4章についてのまとめ

本章では、メタ行動を共有地の悲劇に導入し、そのメタ行動の最適化を行う事で、共有地の悲劇における悲劇の回避を示す事が出来た。即ち、確定的なルールを用いて記述された環境において、メタ行動の自律的な獲得と、メタ行動の導入による問題解決アプローチの有効性を確認することが出来た。

本章以降では、より不確実性の高い問題である造作物構築行動の獲得に対してメタ行動の導入を行う。その為に、メタ行動である造作物構築行動の自律的な獲得に関する実験を行った。

## 5. 2次元シミュレーションにおける仮想生物

本章以降では、メタ行動の現実的な問題への適用についての議論を行っていく。実際的な生物におけるメタ行動とは何かを考えた場合、巣のような造作物の構築行動の進化的な獲得がその一例としてあげられる。本章では、その様な事例におけるメタ行動の導入として、2次元シミュレーション上における仮想生物の構築物構築行動の獲得に関して議論を行う。

### 5.1 問題の性質について

造作物構築行動のシミュレーションの説明を行う前に、この問題で言う所の複雑さというものについて議論を行う。即ち、ゲーム理論の問題の1つである共有地の悲劇と造作物構築行動のような問題との間にどのような違いがあるのかについて説明を行う必要があると考えられる。

ここで言及されている複雑さとは、行動のタイミングとその報酬が与えられるタイミングとの間が一对一になっていない事を示している。図 5.1 に行動とその報酬の関係を示す。(a)は第4章で取り扱った拡張版共有地の悲劇の場合、(b)は造作物構築行動などの場合の関係を示す。拡張版共有地の悲劇では、ある行動の直後にその評価が与えられる為、その行動の評価が安定的に与えられる。しかし、実際的な環境では、ある行動のシーケンスに対して報酬が与えられる事が多く、適切な行動シーケンスを獲得することが重要になってくる。即ち、本研究では、間接行動の効果が出るまでの時間的な遅れ、適切な行動シーケンスを獲得する為の問題空間の広さ、外乱の発生しやすさを指して、問題がより複雑であると言及している。

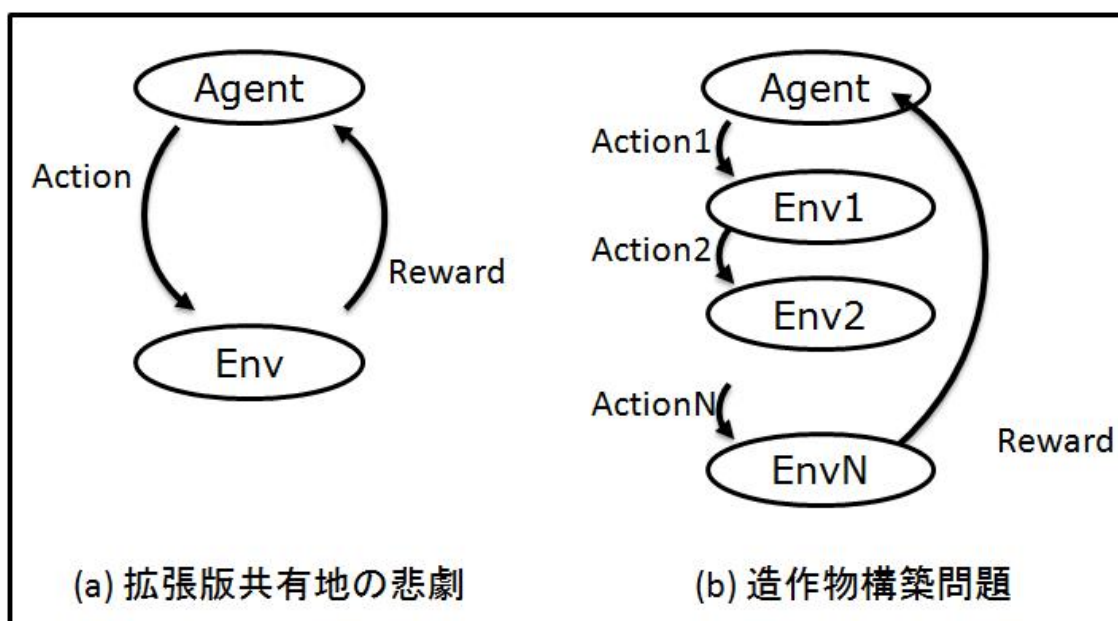


図 5.1 問題の特徴の違い

## 5.2 自然生物におけるメタ行動

自然の生物の構築する構造物の中でも、もっとも一般的なものは巣であると考えられる。ビーバー、アリジゴク、アリ、ハチ、クモなど、多くの生物はある機能性を持った巣を構築することで、環境を自身の生存に有利な状態へと変化させる事ができる。それらの機能性は大別すると、防護、住居、捕獲などが考えられる。その中でも、本章以降では、捕獲用の構造物の構築に焦点を絞り、研究を進める。

実際の生物における構造物構築行動の観察結果などを考慮し、環境設定を決定した。本研究では特に、クモの生態について焦点を絞った。

### 5.3 クモの生態について

#### 5.3.1 身体的特徴について

##### (1) 身体構造について

図 5.2 と図 5.3 はクモの身体構造に関するものである。昆虫とは異なり、前体部と腹部の 2 部分から構成されている。触肢 1 対、脚 4 対、8 個の眼と 1～5 個の糸疣を持つ。全身には体毛が生えており、この体毛によって風や振動、接触を感知することが出来る。また脚先には爪が付いており、接触センサの役割を果たしている。脚の節には、琴状器官と呼ばれる振動検知を行う器官があり、巣上において振動源の方向を検知することが出来る。

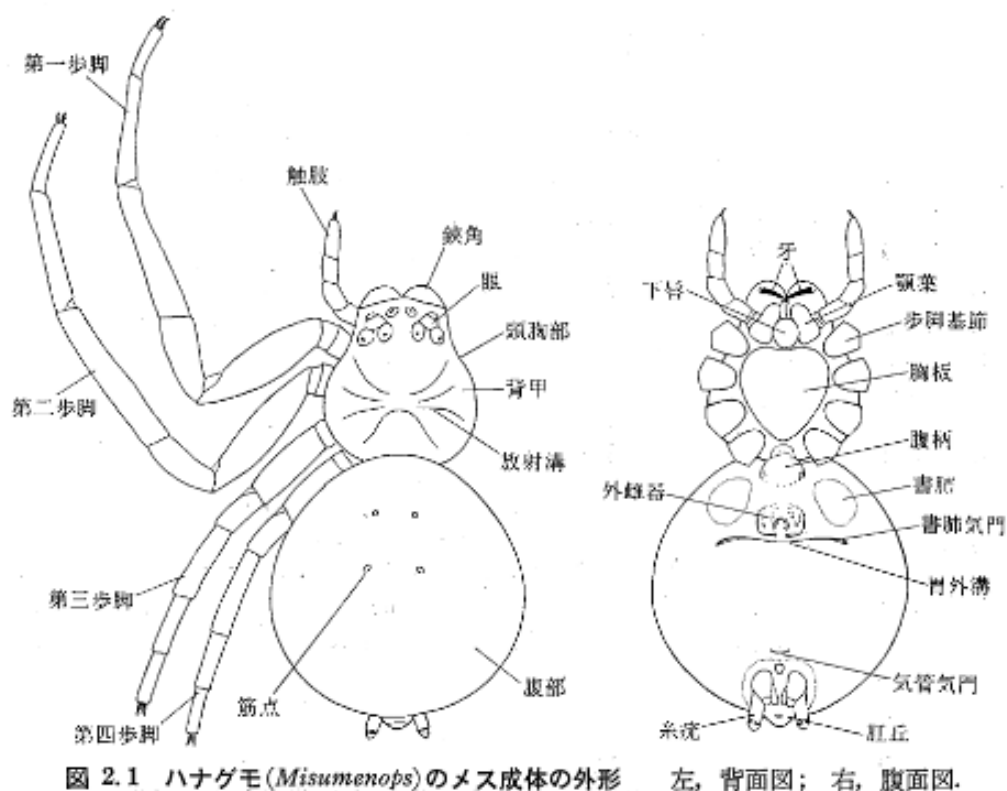


図 5.2 クモの身体構造の例

※図 5.2, 5.3 は参考文献[8]からの引用である



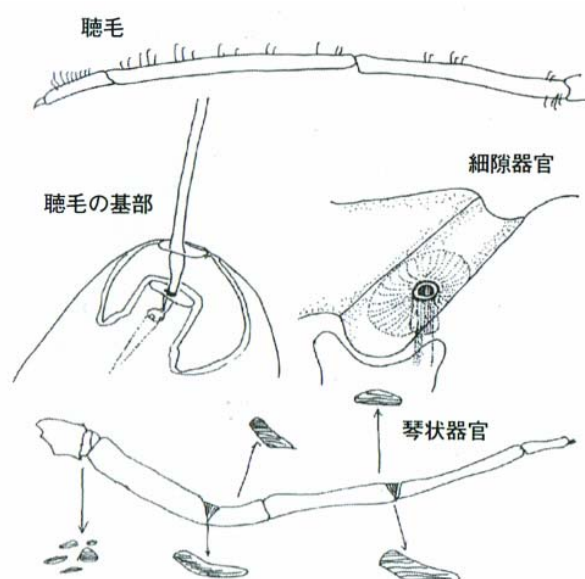


図 5.3 クモの感覚器官の例

## (2) 視覚について

クモの場合、2つの主眼と6つの副眼を持っている。昆虫とは異なり、いずれの眼も単眼構造になっている。眼の配置は種によって異なり、造網生活を行う種の場合、より散開している傾向がある。

まず、昆虫の眼やクモの眼に共通する特徴として、運動視のみが可能である事が挙げられる。即ち、静止しているものを視認することが出来ず、動いているものしか視認できない。主眼は最も前側に位置する眼で、主に近距離の視認に用いられ、3~4種類の色（青、緑、黄、紫外線）を認識することができる。副眼は、遠距離の視認に対応しており、緑色のみを認識できる。また、光に対する感知速度が速い。遠くにいる獲物を副眼で見ながら接近し、近くにくると主眼で見ながら行動を行うといった役割分担が見られる。

ちなみに、昆虫の視覚は、2個の複眼と2~3個の単眼で構成されている。複眼は1ブロックに付き、1画素を認識していると考えられる。イエバエは12000

画素、トンボの場合は 50000 画素と言われている。これらは比較的粗い画像であるが、視野角が広いという特徴がある。以上のように、空間解像度は人間に比べて低くなっているが、時間解像度は人間よりもはるかに高く、高速アクロバット飛行などに貢献している。それに対して、単眼は明暗の検知をする単純な光受容器であり、解像度が低い代わりに、明暗の検知に優れている。昆虫は、複数個の単眼を用いる事によって、太陽の位置や水平線の傾きを検出することが可能であると考えられており、それらによって自身のロー、ピッチ、ヨーの傾きを検出し、飛行制御に役立てていると考えられている。

### (3) 記憶能力について

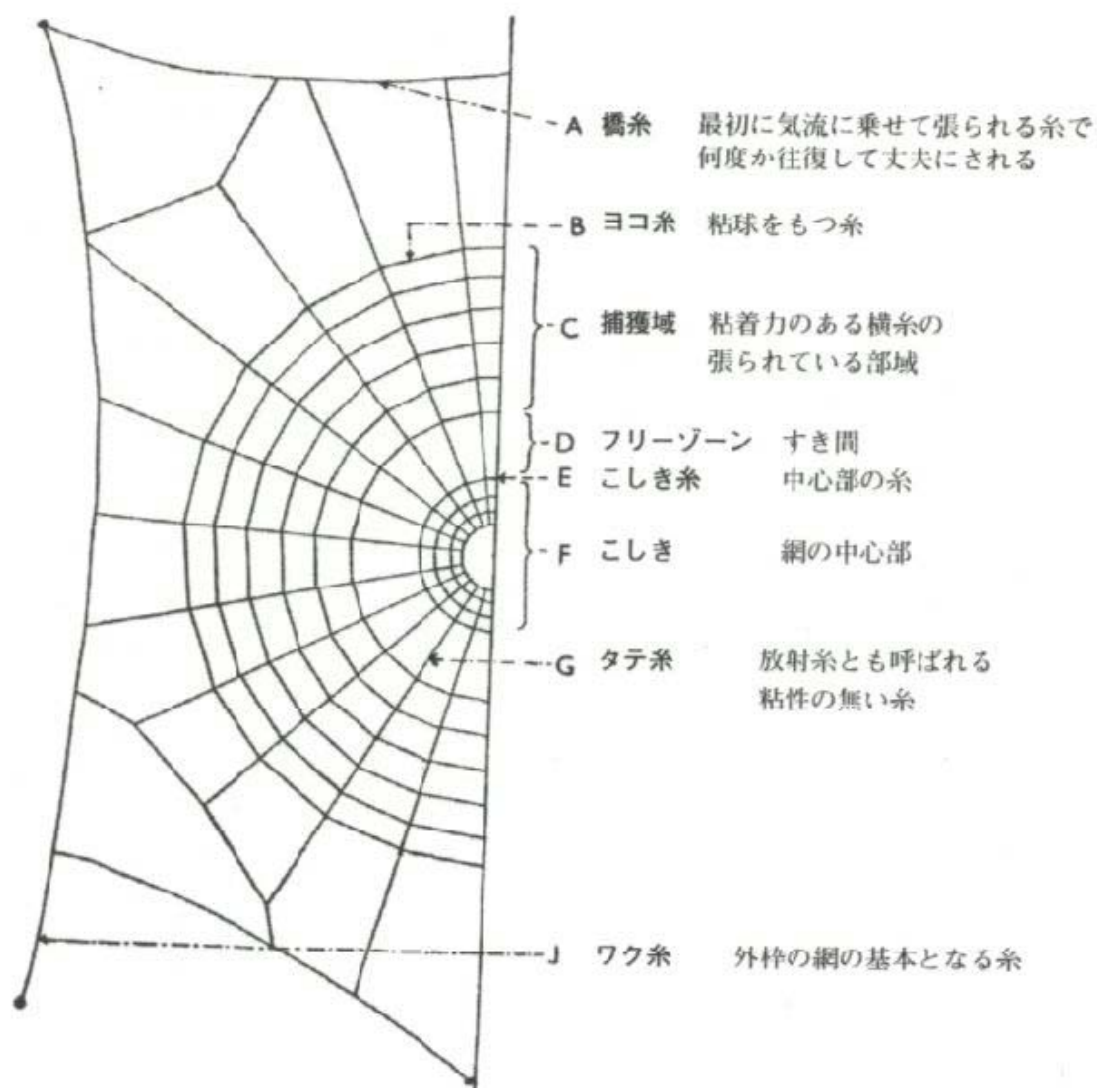
昆虫もクモもある程度の短期記憶の保持が可能であることが、実験的に検証されている。巣を持つ昆虫の殆どは巣の周囲の風景を記憶可能であり、特に縄張りや決まった飛行ルートを持つ昆虫では場所記憶の能力が特に発達している。ミツバチの場合は、その行動圏内で、いつ、どこで、どのような花を訪れば蜜が得られるのか、花の形、匂い、色、場所、開花時間を学習可能であるし、餌場と巣の間の距離や方向を記憶可能である。昆虫やクモの記憶は、脳神経節内のキノコ体と呼ばれている部分で行われている事が実験的に確認されている。しかし、その詳細については未だ解明されていない。

#### 5.3.2 造網行動について

クモの巣の中で一般的なイメージとして扱われているものは、円網、とくにらせん網であるだろう。らせん網は少量の糸で広い空間を占有することが出来、頑強でありながら、張り替えや補修が簡易で、垂直や水平、斜めにも掛ける事ができるという特徴がある。図 5.4 に一般的ならせん網の構造と部位の説明を示

す。

網の半分を示し、網の名称を一度、確認しておきます。



(T. H. Savory の「The Spider's Web」をもとに一部改変して引用)

図 5.4 クモの巣の構造

※参考文献[8]からの引用

一般的な巣の張り方は大まかには以下の通りである。前の行動で作られた巣の部位が次に作られる部位を規定しており、行動は基本的には後戻りしない事が、造網行動の特徴として挙げられる。

1. 橋糸を張り、補強する。張り方は様々で、風に糸をなびかせたり、歩いて移動したりなど環境に依存する。
2. 巣を三叉の状態にし、中心（こしき）を決定する。
3. こしきから放射糸を枠に向けて張り、張った糸にそって中心へと戻る。こしきに戻った際、張力の確認が行われ、張力が不十分であれば次の放射糸を張り、十分であれば足場糸を張るプロセスへと移行する。
4. 足場糸を広めの間隔でこしきから外側へとらせん状に張っていく。足場糸の間隔はクモの脚の長さによって決定される。
5. 外側からこしきに向かって、粘着糸をらせん状に張っていく。その際、足場糸を辿りながら移動を行い、粘着糸を張りながら、用済みの足場糸を切って回収していく。

また、クモの造網行動についての特徴をまとめると、以下の点が挙げられる

- ・捕食行動や造網行動は生得的なものであり、学習によって大幅に変わる事はない。また、種によって異なる。
- ・前に作られたものが次に作られるものを規定している。行動は連鎖的であり、基本的にやり直しは行われない。大きな環境の変化があったときのみやり直しが行われる。
- ・造網行動時には、視覚情報は用いられず、脚先の爪による接触センサ、体毛による振動センサ、脚の節にある琴状器官などを用いている。

### 5.3.3 補虫行動について

補虫行動についても、造網行動と同様に、特定の行動シーケンスが遺伝的に決定されている可能性が指摘されている。即ち、ある行動の結果によって作られた変化が次の行動を決定するという反応と行動の連鎖によって振る舞いが形成されている可能性である。

### 5.3.4 クモの生態を考慮したシミュレーション上のエージェントの設定

上記のクモの生態のまとめより、以下の設定を考慮した

- ・ 生得的な造作物構築行動を表現する為に、遺伝的アルゴリズムと人工ニューラルネットワークの組み合わせ手法を用いる。これは、**Evolutionary Robotics** で主に用いられている手法と同様である。
- ・ センシングする内容について、主に近接センサのようにきわめて近距離的なもののみを用いる。内部状態については、現在位置と向き、残りの巣の材料の感知が可能とした。

## 5.4 シミュレーションの概要

このシミュレーションでは、構造物の構築を行う捕獲者エージェントと捕獲対象である獲物エージェントの2種類のエージェントが存在している。図 5.5 にシミュレーションの大まかな流れについて示す。流れとしては、Virtual Creature や Evolutionary Robotics における一般的なシミュレーションと同様、遺伝的アルゴリズムのシミュレーションが基になっている。即ち、遺伝子群を生成し、各遺伝子について試行を行い評価し、その評価をもとに次世代を生成する。試行の際には、各遺伝子はエージェントの意思決定機構である人工ニューラルネットワークの荷重として変換される。どのような試行を行うかは実験設定によって異なるためここでは割愛するが、基本的には捕獲者エージェントが造作物の構築を行い、獲物エージェントは環境内の移動を行う。

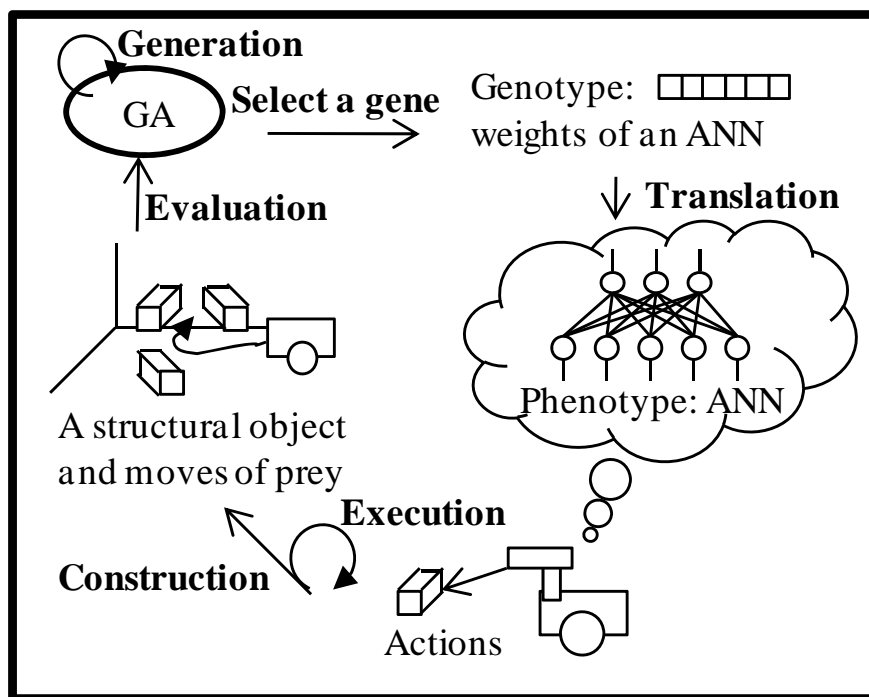


図 5.5 シミュレーション概要

## 5.5 造作物構築行動の獲得のメタ行動の導入の枠組みによる表現

ここで、捕獲用造作物の構築行動の獲得をメタ行動の導入の枠組みで表現したものを図 5.6 と図 5.7 に示す。図 5.6 は、造作物構築行動の進化的獲得に関する実験である実験 1 を示している。実験 1 では、メタ行動である造作物構築行動の獲得のみに焦点を当てている。その為、捕獲者のみが造作物構築行動を行い、その後獲物が環境に流入してくる、いわゆる交代制のシミュレーション設定となっている。実験 2 では、メタ行動と直接行動である直接捕獲の両方を獲得可能な場合について検証を行っている。その為、捕獲者が直接捕獲と間接捕獲を同時に行う事のできる環境設定が必要となる。そこで、実験 2 では、捕獲者と獲物が並列的に動作可能であるような設定を採用している。即ち、捕獲者の行動途中でも、一定時間毎に 1 体の獲物が環境に流入してくる、並列的なシミュレーション設定とした。

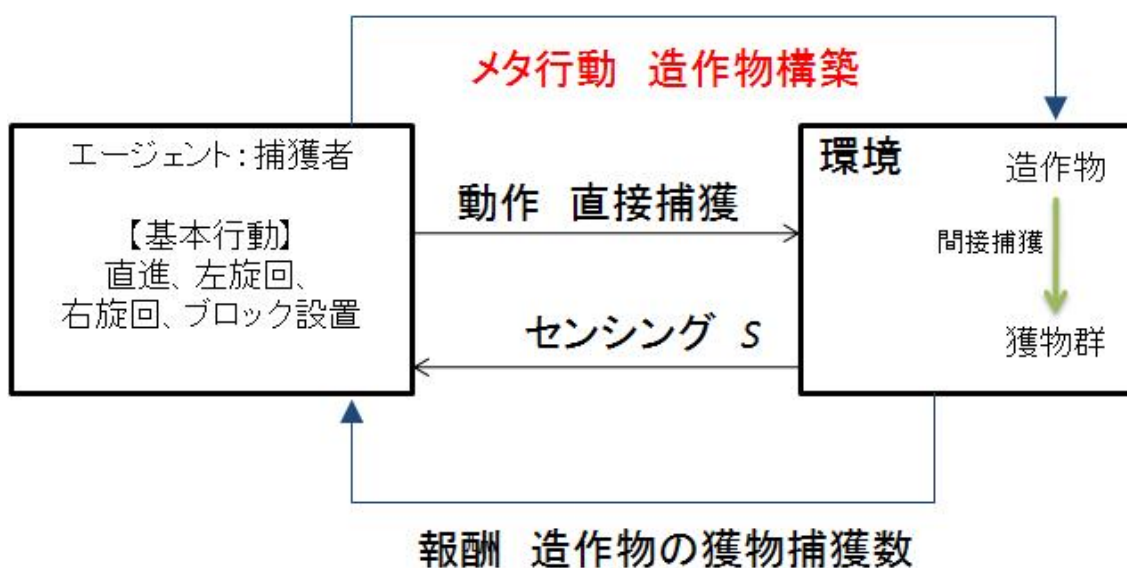


図 5.6 造作物構築行動のみの獲得を行う場合のメタ行動の枠組みによる表現

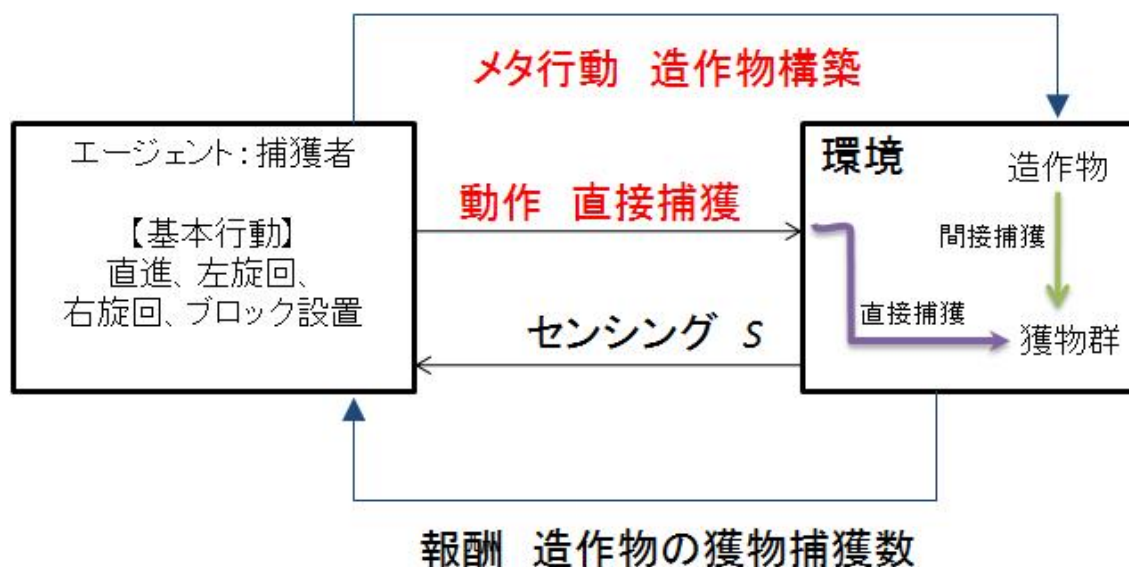


図 5.7 直接捕獲と造作物構築行動の両方の獲得を行う場合のメタ行動の枠組みによる表現

## 5.6 実験 1 造作物構築行動の進化的獲得

本章のシミュレーション環境は、実験 1 で取り扱うものと実験 2 で取り扱うものの 2 種類がある。実験 1 のシミュレーション環境では、捕獲者エージェントと獲物エージェントは同時に存在することではなく、順番に行動を行う。実験 2 では、捕獲者エージェントと獲物エージェントが同時に行動を開始する環境を採用している。

図 5.8 に実験 1 のシミュレーション環境の詳細を示す。環境は、2 次元格子状であり、エージェントやブロックなどの要素は 1 タイルの大きさである。シミュレーションエリアの一端には、獲物エージェントの環境への流入口であるスタートエリアがある（図内右端の青色）。その反対側には、獲物エージェントの環境からの流出口であるエスケープエリアがある（図内左端の灰色）。シミュレ



ーションの流れとしては、まず捕獲者エージェントが環境中心部に出現し、構造物構築行動を行う。一定時間経過後、捕獲者エージェントは環境から取り除かれ、代わりにスタートエリアに全ての獲物エージェントが一斉に出現する。獲物エージェントはエスケープエリアに向けて移動を開始する。全ての獲物エージェントの移動が終了すると、1 試行が終了する。この実験設定においては、捕獲者エージェントの目的は、造作物で獲物エージェントをスタックさせることである。獲物エージェントの目的は、脱出エリアへと到着することである。

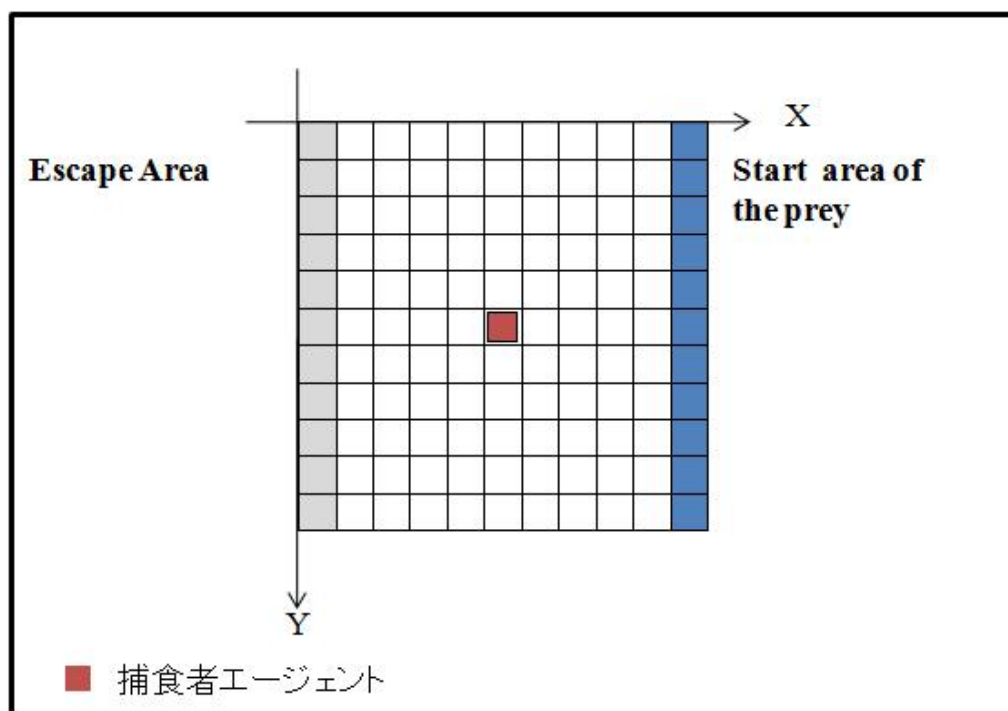


図 5.8 2次元シミュレーションにおける環境の概要

### 5.5.1 獲物エージェント

獲物エージェントは、いわゆる単純反射エージェントの一種である。本章の実験では、獲物エージェントの行動規則は2種類用意されている。1つは前進のみの行動規則であり、もうひとつは簡単な障害物回避の行動規則である。どち

らの行動規則が採用されるかは、実験設定による。

前進のみを行う獲物エージェントの場合、進行方向にブロックが存在すると、スタック状態となり、以降獲物エージェントは行動が出来なくなる。それに対して、障害物回避行動を持つ獲物エージェントの場合は、進行方向にあるブロックを回避して行動することが出来る。その障害物回避ルールは以下の通りである（図 5.9）。

- ・ 進行方向にブロックがない場合、獲物エージェントは直進する
- ・ 進行方向にブロックがない場合、進行方向に対して斜め前の左右どちらかをランダムに選んで移動する。移動後の向きは、移動前の進行方向と同じとする。

即ち、進行方向の斜め左前、前、斜め右前の 3 方向にブロックがある場合は、スタック状態となり、移動できなくなる。これは、造作物の構築コストが上昇している事を示す。

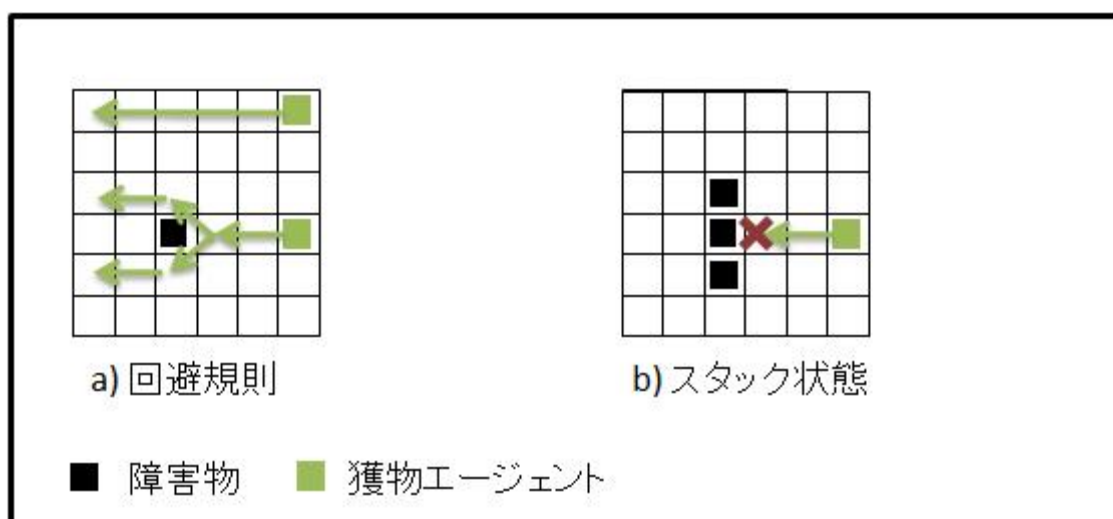


図 5.9 獲物の障害物行動のルール

### 5.5.2 捕獲者エージェント

捕獲者エージェントは、獲物を出来るだけ多く捕獲する事を目的としたエージェントである。特に、この実験 1 では、造作物を構築する事で獲物を出来るだけ多く捕獲することを目的としている。

#### 5.5.2.1. センシング

捕獲者エージェントは、周囲の 1 近傍、計 8 か所の環境情報をセンシング可能である。具体的には、1 タイルにつき、ブロックの有無をセンシングする。センシングするタイルの順番は、捕獲者エージェントの向きに応じて相対的な順番となっており、捕獲者エージェントの正面からみて左上を一番とし、右下を最後としている。図 5.10 に、その順番について図示したものを示す。環境情報以外にも、捕獲者エージェントは内部状態として、現在位置、現在の正面方向、設置できるブロックの残量を持っている。ちなみに、捕獲者エージェントの向きは 8 方向の 45 度刻みで設定されている。捕獲者エージェントは、これらの環境からのセンシング結果と、内部状態を意思決定機構への入力として扱う。

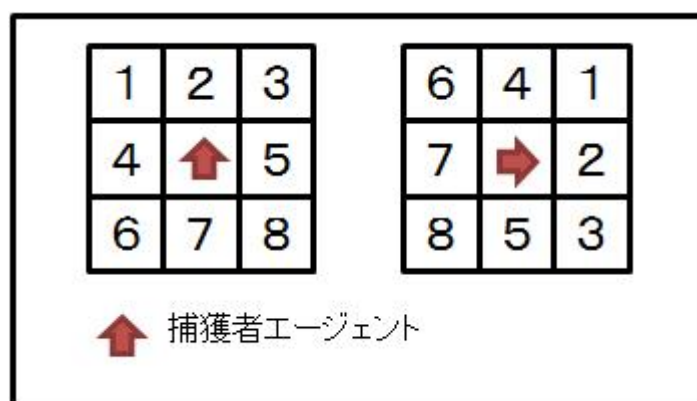


図 5.10 周囲のブロックの有無のセンシングの相対的な順番

#### 5.5.2.2. 基本的な行動

捕獲者エージェントの基本的な行動は、前進、左旋回、右旋回、ブロックの設置の 4 種類である。前進は、進行方向に対して 1 ブロック分移動を行う。図 5.10 中の例で言うと、2 番のタイルに移動を行う事になる。但し、移動先にブロックがある場合は、何も行動できないと判断され、行動が無効となる。左旋回は、エージェントから見て左側へ 45 度回転、右旋回は右側へ 45 度回転を行う。ブロックの設置は、エージェントの進行方向 1 タイルにブロックの設置を行う。ただし、既にタイルの上にブロックが設置済みの場合、行動は無効となる。

捕獲者エージェントは、上記のセンシングデータと内部状態を入力として意思決定を行い、上記の基本的な行動のいずれかを選択する。意思決定の詳細については、次のセクションにて述べる。

#### 5.5.2.3. 意思決定

捕獲者エージェントは、2 層ニューラルネットワーク (2LANN) を意思決定機構として用いている。このニューラルネットワークの荷重は遺伝子から翻訳された生得的なものであり、学習による変化は考慮されていない。入力としては、周囲の環境情報と内部状態を用いる。図 5.11 は意思決定に関わる入力、出力とニューラルネットワークの関係を図示したものである。入力は、全て $[-1.0, 1.0]$ の範囲に線形変換される。環境情報は、検知する周囲の環境一ブロックにつき 1 入力として扱い、ブロックがあれば 1.0、なければ-1.0 の値がセンシング結果となる。出力値は、 $[0.0, 1.0]$ の範囲となっている。出力ユニットは各行動に対して割り当てられており、最も出力値の高い出力ユニットの行動が選択され

る。入出力関係をまとめると以下のようなになる。

入力情報 = {環境情報、内部情報、閾値荷重}

環境情報 = {-1.0 (ブロック無し) , 1.0 (ブロック有り)}

内部情報 = {角度、現在位置 X 座標、現在位置 Y 座標、設置可能残りブロック数}

出力情報 =  $\{O_j = [0.0, 1.0] \mid 0 < j < Num_{out}\}$

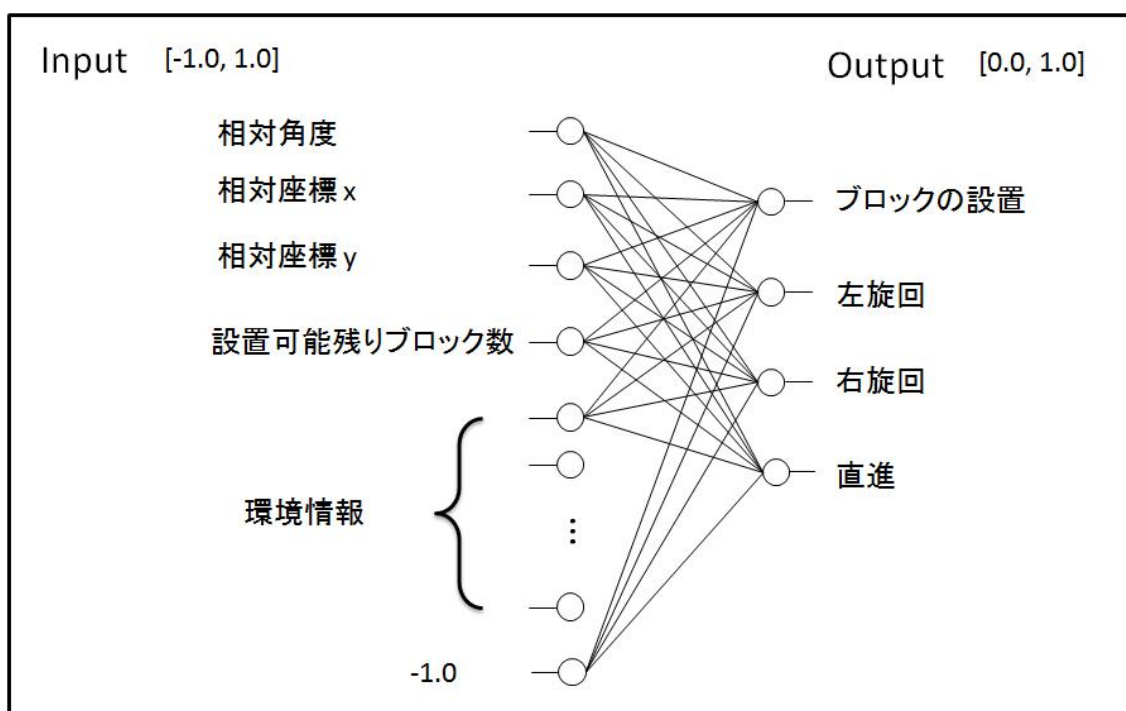


図 5.11 意思決定機構

### 5.5.3 評価関数

実際の生物による造作物構築行動の進化的な獲得を考えた場合、その評価は、基本的には生き残れるかどうかという事になってくる。生き残る為には、より効率的に多くの獲物を捕食する必要がある。即ち、作られた構造物に対する構造的な評価ではなく、あくまでも生存に関する機能のみが評価されていると言える。その点を考慮して、本研究でも、基本的には、造作物の獲物の捕獲機能について評価を行う。その為に、評価関数では、ブロックで足止めする事の出来た獲物の数  $N_{prey}$  を評価として扱い、造作物の評価とする。評価関数を以下に示す。

$$Fitness_1 = \alpha N_{prey} \quad (5.1)$$

但し  $\alpha$  は定数

但し、構造物の構造に関する評価を行った場合に、どのような影響が造作物構築行動の獲得に対して現れるか検証する為に、構造の評価を行う場合の評価関数を用いた実験も行う。その評価関数では、構造物の Y 軸方向での長さの評価を行っており、Y 軸方向に直線的に長い程、評価が高くなる。式は以下の通り。

$$Fitness_2 = \alpha \cdot N_{prey} + \beta \cdot StEval(S_i) \quad (5.2)$$

但し、 $\alpha$ 、 $\beta$  は定数、 $StEval(S_i)$  は造作物  $S_i$  の構造の評価を行う関数であり、以下の式で表される。

$$StEval(S_i) = \begin{cases} 0 & (Len_{Block} < 3) \\ Len_{Block} - 2 & (Len_{Block} \geq 3) \end{cases} \quad (5.3)$$

$Len_{Block}$ は造作物  $S_i$ において Y 軸方向に連続して並んだブロックの数であり、 $StEval$ は Y 軸方向にブロックが 3 個以上並んだ場合は 3 個目以上の並んだブロックの数を評価とし、3 個以下の場合は 0 を評価とする関数である。

#### 5.5.4 実験 1 の設定

2 次元環境における造作物構築行動の進化的獲得を目的として、上述の設定により実験を行った。また、意思決定機構として、2 層 ANN 以外に NEAT のどちらが問題に適しているか検討を行う為に、それぞれのニューラルネットワークを用いた場合の結果の比較を行った。

実験パラメータを以下に示す。

##### 環境

- ・ 環境の広さ                       $19 \times 19 = 361$  タイル（移動可能部分のみ）
- ・ 獲物エージェントの数              19 体
- ・ 最大ブロック数                      19 個

##### 評価関数

- ・  $\alpha$                       10
- ・  $\beta$                       10

#### 遺伝的アルゴリズム関係

- ・ 突然変異率                    0.10
- ・ 交叉率                        0.70
- ・ 保存エリート                1 体
- ・ 1 世代の個体数            50 体

#### 5.5.5 実験結果 1

##### 5.5.5.1. 評価関数 1 を用いた場合の実験

図 5.12 に評価関数 1 を用いた場合の結果を示す。図は各世代のエリート個体の評価値の遷移を表している。

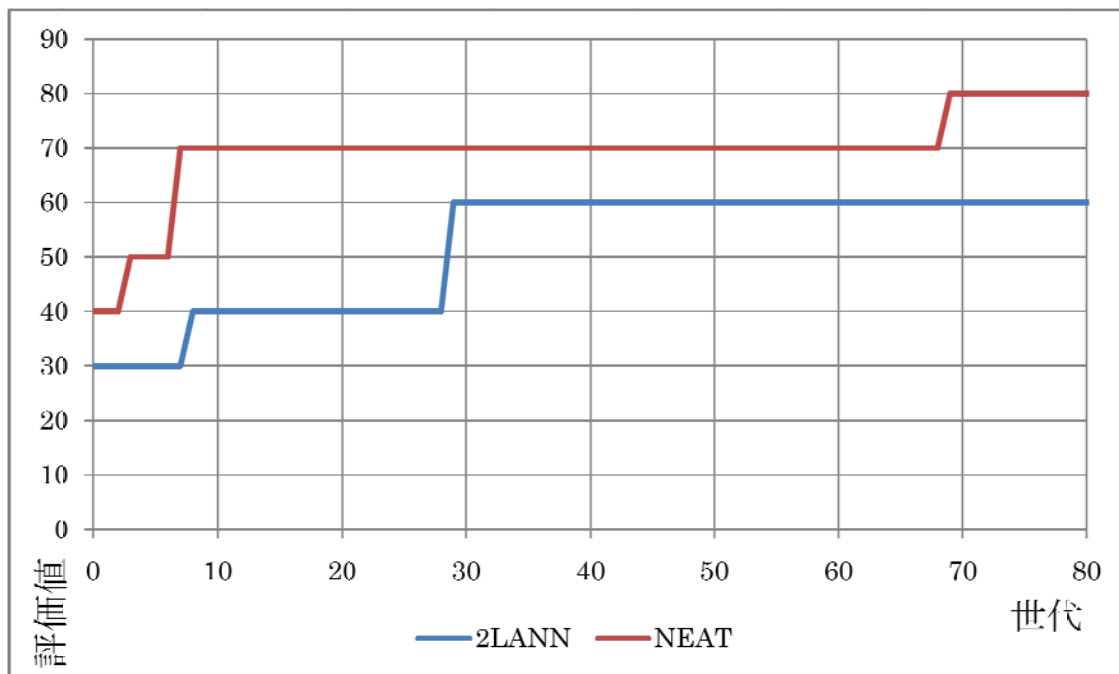


図 5.12 評価関数 1 の場合の評価値の遷移

NEAT を用いた場合、2LANN を用いた場合よりも速く収束している事がわ



かる。しかし、最終的な収束値としては、大きな違いが見られない事がわかる。また、評価関数の違いによって行動の大きな変化は見られなかった。獲得された行動は次のような特徴があった。初期頃の世代では、その場で回転を行い、周囲にブロックを配置する行動を獲得していた。その後、配置したブロックの周りを移動しつつ、ブロックを配置する行動が獲得された。実験毎に異なる行動が獲得されており、ジグザグで斜めに進んでいく場合とらせん状に回転しながらブロックを配置していく場合があった。図 5.13 と図 5.14 に、獲得された造作物の例を示す。図中「<」は獲物、「□」はブロック、「E」は脱出エリア、矢印は捕獲者を示す。しかし、この実験環境で最も効率の良いと考えられる、Y 軸方向に対して直線的な構造を持つ造作物の獲得には至らなかった。

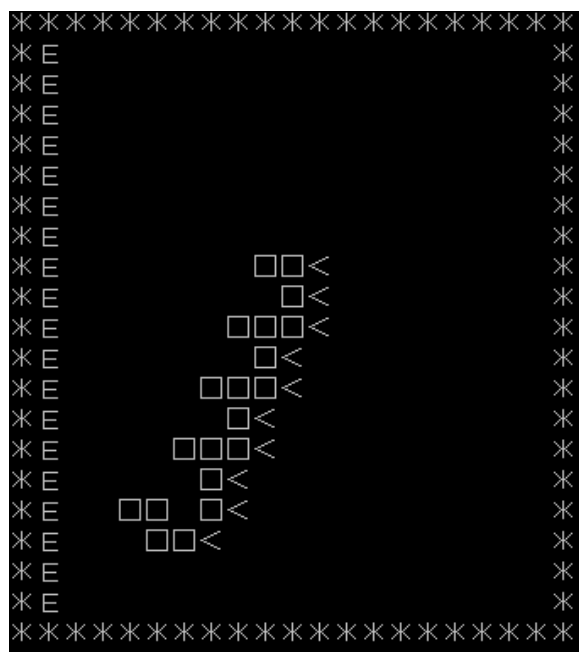


図 5.13 評価関数 1 の場合の造作物の例 1

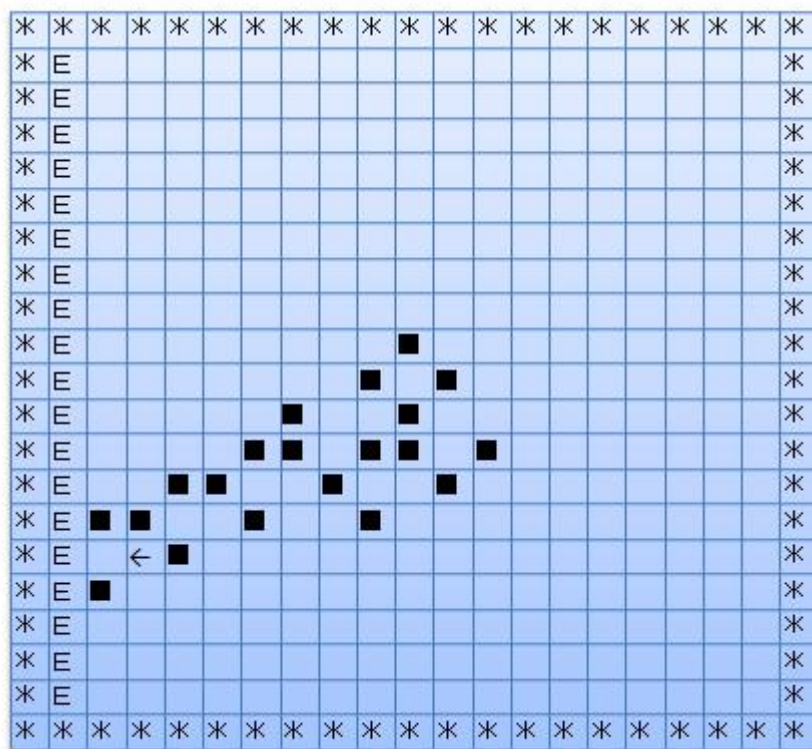


図 5.14 評価関数 1 の場合の造作物の例 2

#### 5.5.5.2. 評価関数 2 を用いた場合の実験

図 5.15 に、評価関数 2 を用いた場合の結果を示す。評価関数 1 を用いた場合と同様、図は、エリート個体の評価の遷移を表している。こちらの場合も NEAT 手法の方が 2 層 ANN よりも収束が早い。しかし、収束値としては、ほぼ同等であると言える。

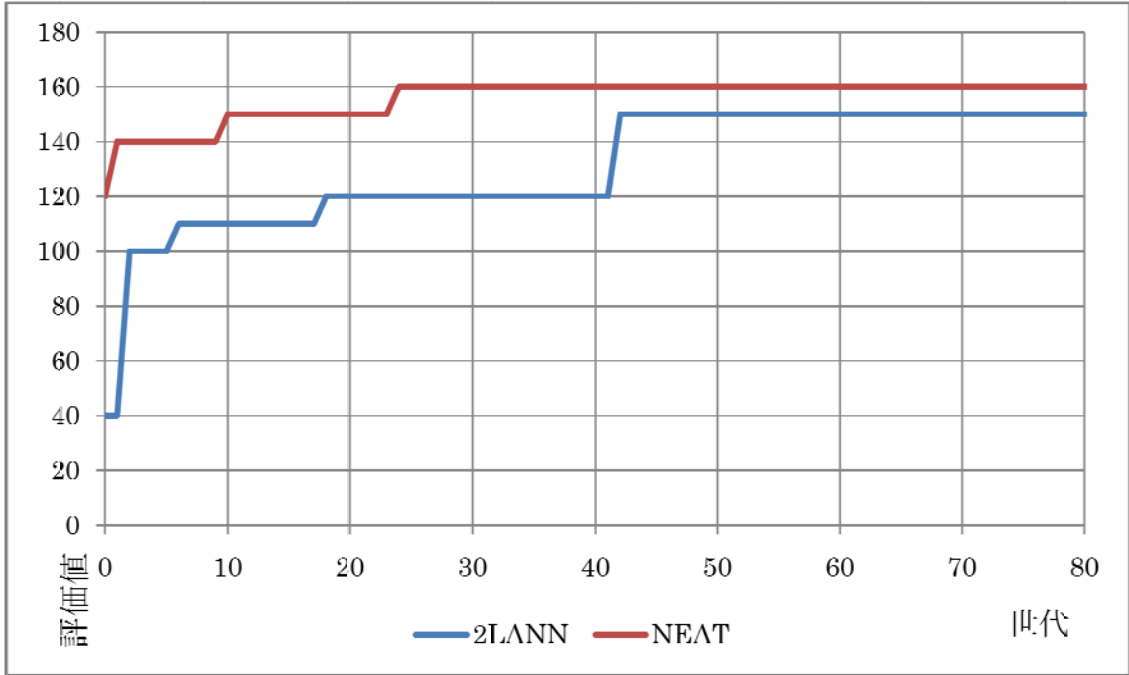


図 5.15 評価関数 2 の場合の評価値の遷移

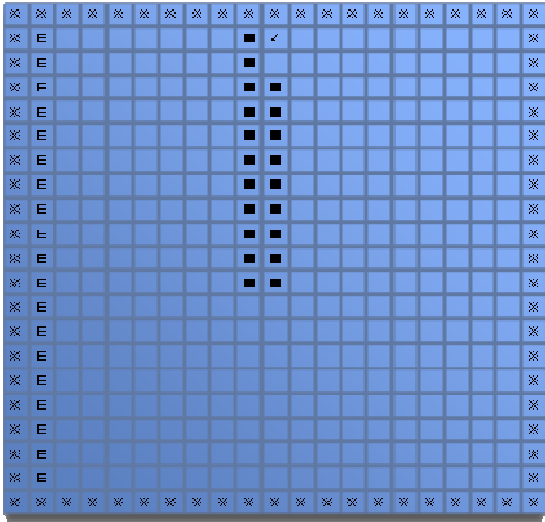


図 5.16 評価関数 2 の場合に獲得された造作物の例 1

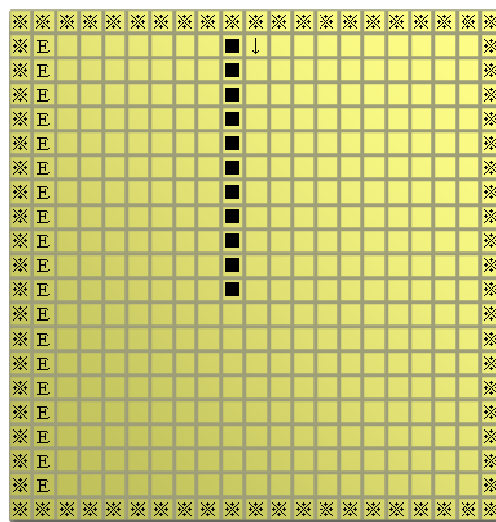


図 5.17 評価関数 2 の場合に獲得された造作物の例 2

図 5.16 と図 5.17 は、獲得された構造物を示している。結果として、評価関数の中で評価されている構造的特徴を持った構造物を構築する行動が獲得されている事がわかる。

#### 5.5.5.3. センサシングの仕方が異なる場合について

センシングの方法による造作物構築行動への影響を検討する為に、環境のセンシングの仕方を変更した場合の実験を行った。これまでの実験では、捕獲者の向きに対して相対的な順序でタイル上のブロックの有無についてセンシングを行っていた。この実験では、捕獲者の向きに関係しない固定的な順序でタイル上のブロックの有無についてセンシングする方法を採用し、それ以外の条件を 5.5.5.1 の実験と同じとした。図 5.18 は、捕獲者の向きに対して固定的な順序でのセンシングを示したものである。結果として、図 5.16、5.17 で示されたような直線性を持つ造作物を構築する行動を獲得できている。

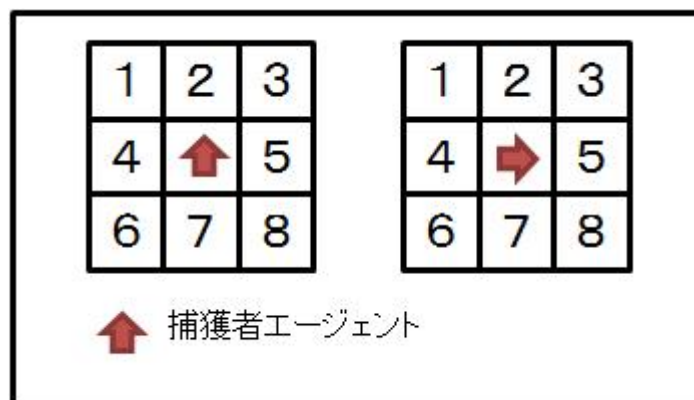


図 5.18 固定的な順序によるブロックの有無のセンシング

#### 5.5.6 実験 1 まとめ

実験 1 を通じて、遺伝的アルゴリズムとニューラルネットワークの組み合わせによる NeuroEvolution 的枠組みを用いる事で、造作物構築行動というメタ行動の一種を進化的に獲得できる事を確認する事ができた。また、用いるニューラルネットワークを何にするかという事についても検討を行った結果、本シミュレーションの場合、NEAT、2LANN のいずれの場合でも同等の結果を得る事が可能であることが分かった。実験 2 以降では、構造が簡単で解析しやすいと考えられる 2LANN を用いて実験を行っていく。

5.5.5.2 節と 5.5.5.3 節で行った実験の結果から、造作物構築行動に影響を与える方法として、評価関数への造作物の構造的な評価の項の追加と、ブロックのセンシングの仕方の二通りを確認する事ができた。しかし、生物の進化を考えると、進化の際に与えられる評価に造作物の評価が直接的に含まれている事は考えにくい。従って、環境や捕獲者の身体、ひいてはセンシングの仕方に造作物構築行動に対して影響を与える要因が含まれていると考えた方が良いと思われる。

## 5.6 実験 2 直接捕獲と間接捕獲の両方を獲得できる捕獲者による実験

間接的行動の導入を考える上で、どのような環境であれば間接的行動を導入できるか、間接的行動を導入する意味があるか、という事を考える事は非常に重要であると言える。第 3 章で述べた通り、間接的行動の獲得は、直接的行動の獲得に比べると獲得しにくい傾向があると予想される。間接的行動は、直接的行動をサポートする事で、直接的行動によって得られる評価の効率を上昇させる事が出来る。しかし、間接的行動を行うという事は、その間に直接的行動による評価の上昇が期待できないという事である。即ち、間接的行動の為の時間、労力などのコストが必要であり、間接的行動を実行している間は直接的行動による評価の上昇がない為、直接的行動のみの場合よりも進化的に獲得されにくい傾向にあると考えられる。

クモを例にとって考えてみる。多くのクモは、捕獲用の罟を作らず、直接追いかけたり、待ち伏せしたりして獲物を直接的に捕獲している。進化を経て、糸を用いて捕獲用の罟を作るような種が出現してきた。罟を作る為には、1 日から 2 日という時間的なコスト、巣の材料であるたんぱく質などの資源的なコスト、巣を構築する際の労力など様々なコストが発生する。巣を作っている間は獲物を捕獲することができないが、巣が完成すれば、自分よりも移動速度が速い獲物、空を飛んでいる獲物など、直接的捕獲では獲得できなかった獲物を捕まえる事が出来るようになる。以上の点を踏まえて考えると、間接的行動による一時的な評価の下降や停滞を上回る評価を後々に獲得できる事、身体的な能力の不足を補える事、これらが間接的行動を獲得する為の重要な要素であると考えられる。

実験 1 では、間接的行動の一種である造作物構築行動の獲得について実験を

行った。実験 2 では、直接的行動と間接的行動の両方を実行可能な捕獲者エージェントの導入を行い、こういった環境設定、条件であれば、間接的行動の獲得が行われるのか、という事について実験を行った。

### 5.6.1 環境設定の変更

直接的捕獲と間接的捕獲を比較する為に、本実験では、環境設定の変更を行っている。実験 1 では、最初に捕獲者のみが環境に出現して行動を行い、一定時間経過後に捕獲者が消失、その後、獲物が環境に一斉に流入して行動する、という流れで実験が行われていた。しかし、直接捕獲を考えた場合には、捕獲者のみが行動する時間は有用ではない。更に、獲物が一斉に現れた場合、同時に複数体の獲物を捕獲する事は難しく、間接的行動の結果と比較するには不相当であると考えられる。そこで、獲物は一定時間ごとにランダムな位置から環境に流入してくる事とした。図 5.19 に変更後の環境を示す。変更としては、(1) 獲物の出現位置と脱出位置の変更、(2) 獲物の出現タイミングの変更の 2 点である。獲物の脱出エリア (図 5.19 灰色部分) を環境の四辺に配置し、その 1 ブロック内側の四辺を獲物の出現エリア (図 5.19 青色部分) としている。獲物は出現位置とは反対の方向に向かって移動を開始する。出現タイミングは、定数によって決定することとした。実験の際には、各世代の各試行で同じ出現位置、タイミングを用いるものとした。

直接捕獲による獲物捕獲の判定は、捕獲者と獲物が同じタイル上に存在している場合としている。捕獲者と獲物の向き、どちらが先にタイル上に存在していたかについては考慮しない。

獲物と捕獲者の身体的な能力を表す為に、移動速度の導入を行った。移動速

度は、一回の行動に必要なステップ数として表される。即ち、「捕獲者0 獲物3」のような表記の場合、捕獲者が毎ステップ行動し、獲物は3 ステップ毎に移動を行う事をしめしている。この移動速度の違いは即ち、直接捕獲の難しさに関係している。捕獲者の移動速度が獲物の移動速度よりも速い場合は獲物を捕まえやすい事が予想されるが、捕獲者の移動速度が遅い場合は獲物を捕まえる事が難しいと予想される。

5.6.5 節では、障害物回避行動を持たない獲物について実験を行い、5.6.6 節では、障害物回避行動を持つ獲物について実験を行った。障害物回避行動の有無は、造作物構築のコストに関連している。即ち、障害物回避行動を持つ獲物を足止めする為には、最低3 つ以上のブロックが必要になる為、造作物構築コストが高くなっている事になる。

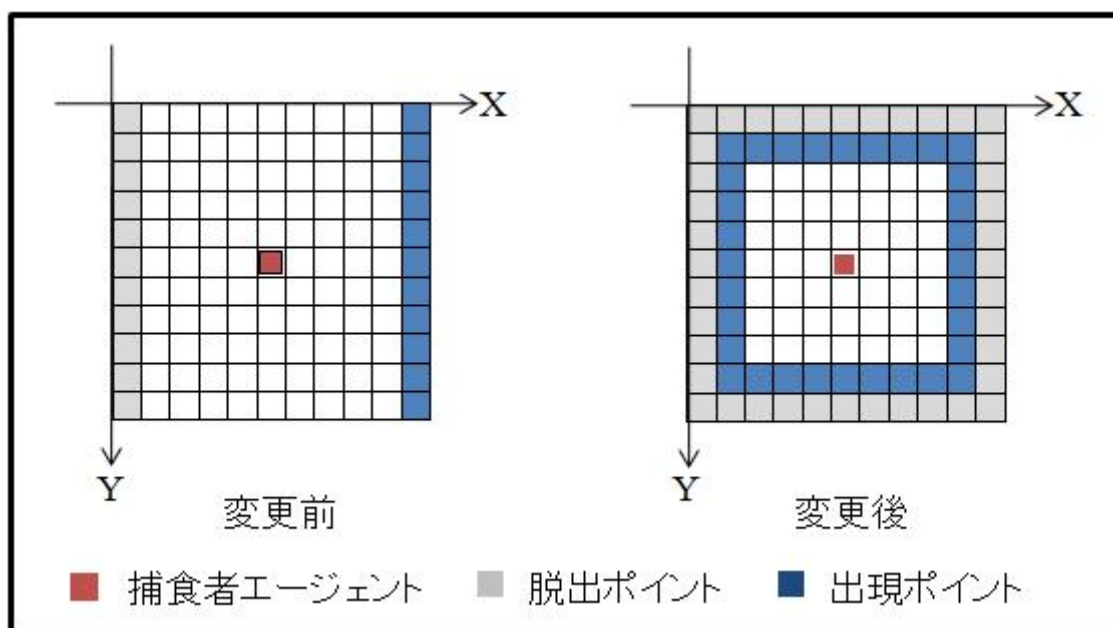


図 5.19 変更後の環境設定



## 5.6.2 直接捕獲用のセンサ

直接捕獲を行う為に、捕獲者には獲物の位置や獲物までの距離を知覚する必要がある。直接捕獲を行うクモの場合、その視野はほぼ 360 度であり、真上以外は見渡す事ができる。その視野内の動く物体の移動速度、物体の大きさを検知し、ある一定の速さ、大きさ、距離であれば、視野の中心に獲物を捕えるように方向を変えながら追跡をすると考えられている。本実験では、これらの特徴を踏まえ、捕獲者は十分な視野を持っており、最も近い獲物までの距離、正面方向の獲物までの距離を検知する事ができるとした。実際には、獲物までの  $x$  軸方向の距離の差分、 $y$  軸方向の距離の差分を検知する最近傍獲物距離センサと、正面方向の獲物までの距離検知する獲物用近接センサを実装し、意思決定機構であるニューラルネットワークへの入力情報として扱っている。図 5.20 に追加したセンサについての図を示す。図 5.21 に直接捕獲を考慮した場合の意思決定機構について示す。

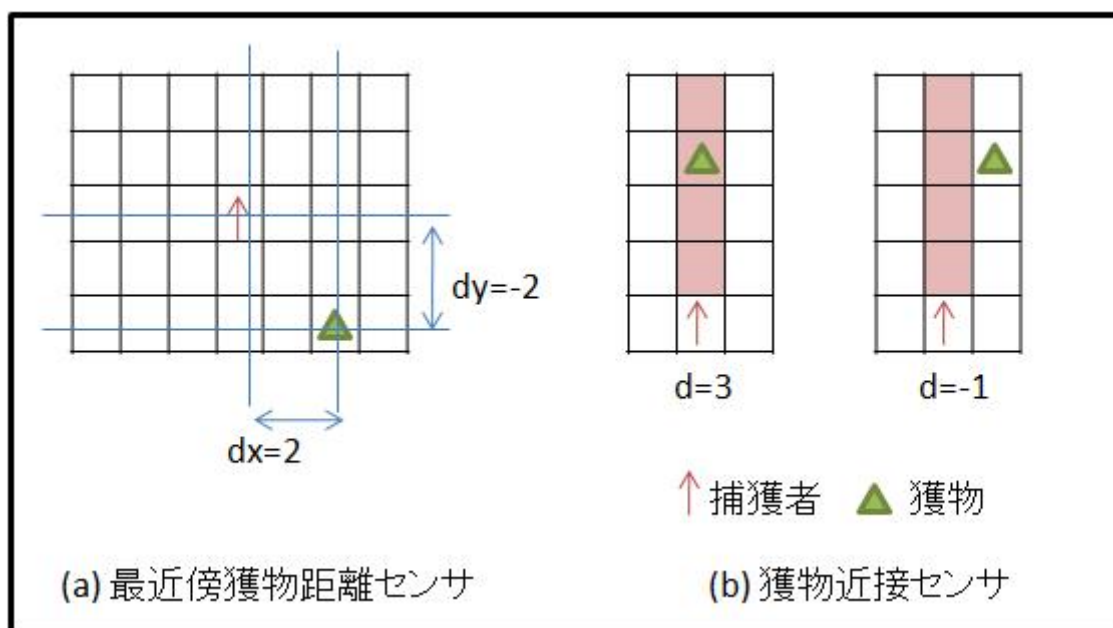


図 5.20 直接捕獲用のセンサ

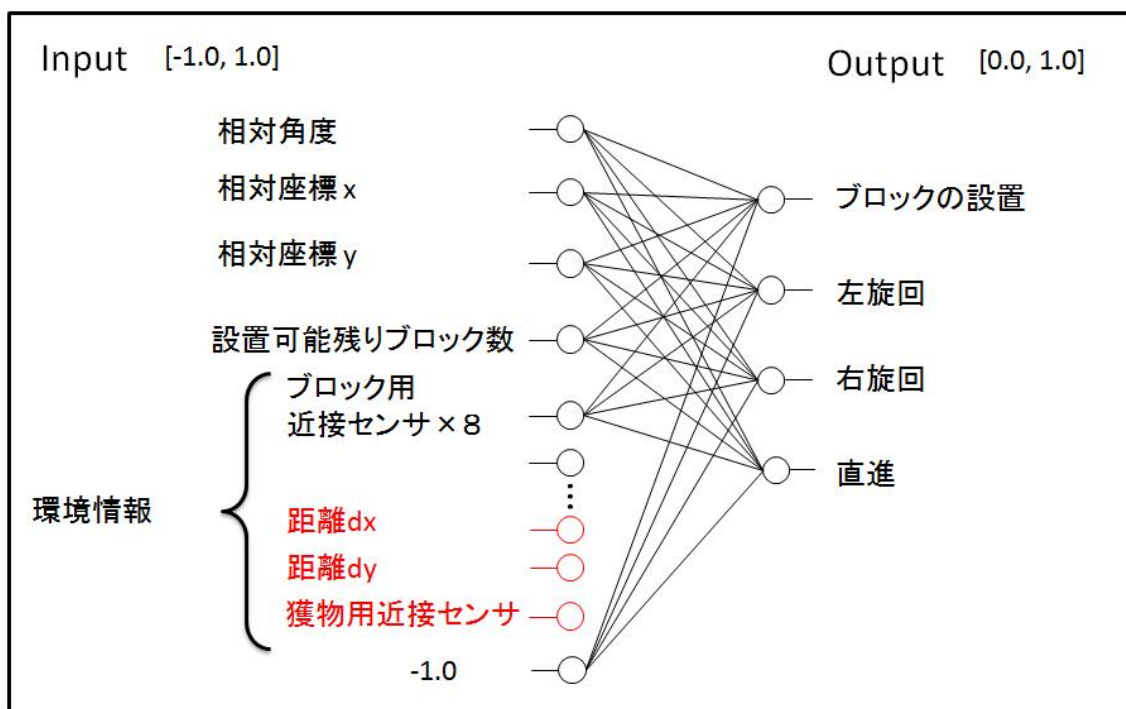


図 5.21 直接捕獲用センサを追加した場合の意思決定機構

### 5.6.3 評価関数

本実験では、直接捕獲による獲物の数  $N_{captured}$  と間接捕獲による獲物の数  $N_{trapped}$  の和を評価値として扱っている。即ち、以下の評価式を採用している。

$$Fitness_3 = \alpha \cdot N_{captured} + \beta \cdot N_{trapped} \quad (5.4)$$

但し、 $\alpha$ 、 $\beta$  は重み係数

#### 5.6.4 実験設定

環境の設定については、上記で述べた変更が反映されたものとなっている。

##### 環境

・ 環境の広さ	19×19=361 タイル (移動可能部分のみ)
・ 獲物エージェントの数	100 体
・ エージェントの出現タイミング	10 ステップ
・ 最大ブロック数	19 個
・ 移動速度	獲物 3 (遅い) < 捕獲者 0 (速い) 獲物 0 (速い) > 捕獲者 3 (遅い) のいずれか 以下のグラフ中では、捕獲者 < 獲物、 捕獲者 > 獲物という表記を用いている
・ 獲物の回避行動	有り 無し のいずれか

その他の部分については、実験 1 と同じ設定になっている。即ち、以下の通りである。

##### 評価関数

・ $\alpha$	10
・ $\beta$	10

##### 遺伝的アルゴリズム関係

・ 突然変異率	0.10
---------	------

- ・ 交叉率                      0.70
- ・ 保存エリート              1 体
- ・ 個体数                      50 体
- ・ 世代数                      100 世代

直接捕獲と間接捕獲に対するコストに関する環境条件について、まとめた表を表 5.1 に示す。

表 5.1 直接捕獲と間接捕獲のコストに関する環境条件

環境条件		捕獲者の獲得する行動		
獲物の回避 行動の有無	移動速度	直接行動のみ	メタ行動のみ	両方
無し	捕獲者 > 獲物	5.6.5.1 節で説明		5.6.5.2 節で 説明
	捕獲者 < 獲物			
有り	捕獲者 > 獲物	5.6.6.1 節で説明		5.6.6.2 で 説明
	捕獲者 < 獲物			

### 5.6.5 実験結果 1 獲物にブロック回避行動がない場合

#### 5.6.5.1. 直接捕獲と間接捕獲を別々に行った結果について

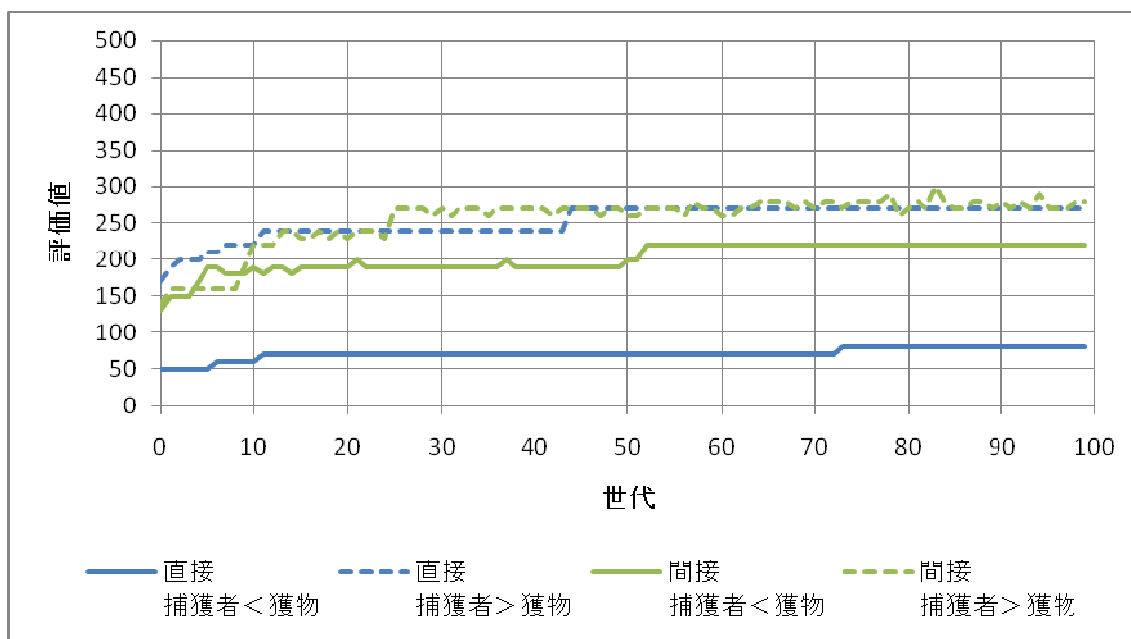


図 5.22 直接捕獲のみと間接捕獲のみでの実験結果

図 5.22 は、直接捕獲のみを行う事ができる捕獲者を用いた場合の結果と、間接捕獲のみを行う事ができる捕獲者を用いた場合の結果を示すものである。通常線は、捕獲者の移動速度が獲物よりも遅い場合を示し、点線は捕獲者の移動速度が獲物よりも速い場合を示す。

直接捕獲の場合、捕獲者の移動速度の違いによって評価値に大きな違いが出ている事がわかる。捕獲者の方が速い場合は最終的には 27 体の獲物を捕獲することが出来ているのに対して、捕獲者が遅い場合は 8 体程度しか捕まえる事が出来ていない。一方、間接捕獲では、獲物の移動速度による違いはそれほど大きくなってはいない。これらの事柄から、移動速度の違いに合わせた進化が行

われている事が確認できた。

#### 5.6.5.2. 直接捕獲と間接捕獲のどちらも選択できる場合

図 5.23 と 5.24 は、図 5.22 で示した直接捕獲のみの結果と間接捕獲のみの結果と、直接捕獲と間接捕獲の両方の場合の結果を示したものである。また、図 5.25 は、直接捕獲数と間接捕獲数の推移を示したものである。移動速度の違いによらず、いずれの場合でも、直接捕獲・間接捕獲の両方を行う捕獲者が最も高い評価値を獲得する事が出来ている。これは、直接捕獲・間接捕獲の両方を行う捕獲者の場合、直接捕獲と間接捕獲を上手く組み合わせることによって、獲物捕獲の効率性を向上させる事が出来ているからだと考えられる。捕獲者の移動速度が速い場合は、直接捕獲と間接捕獲の両方の効率を改善している事が獲物捕獲数の推移から読み取る事ができる。

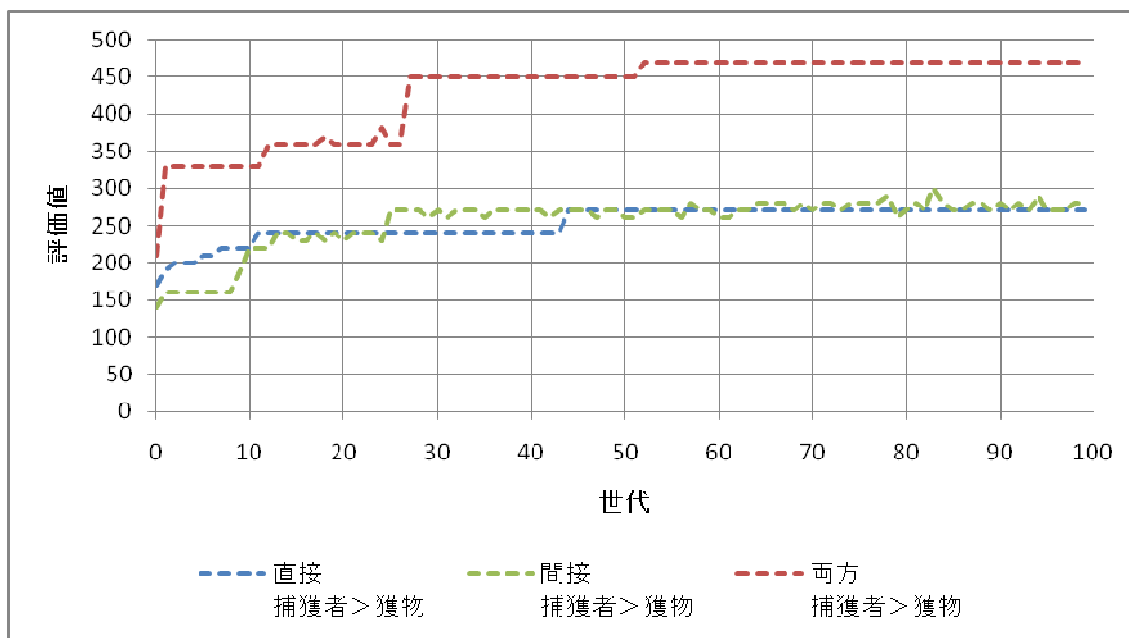


図 5.23 捕獲者の移動速度が速い場合の直接捕獲、間接捕獲、両方の評価値

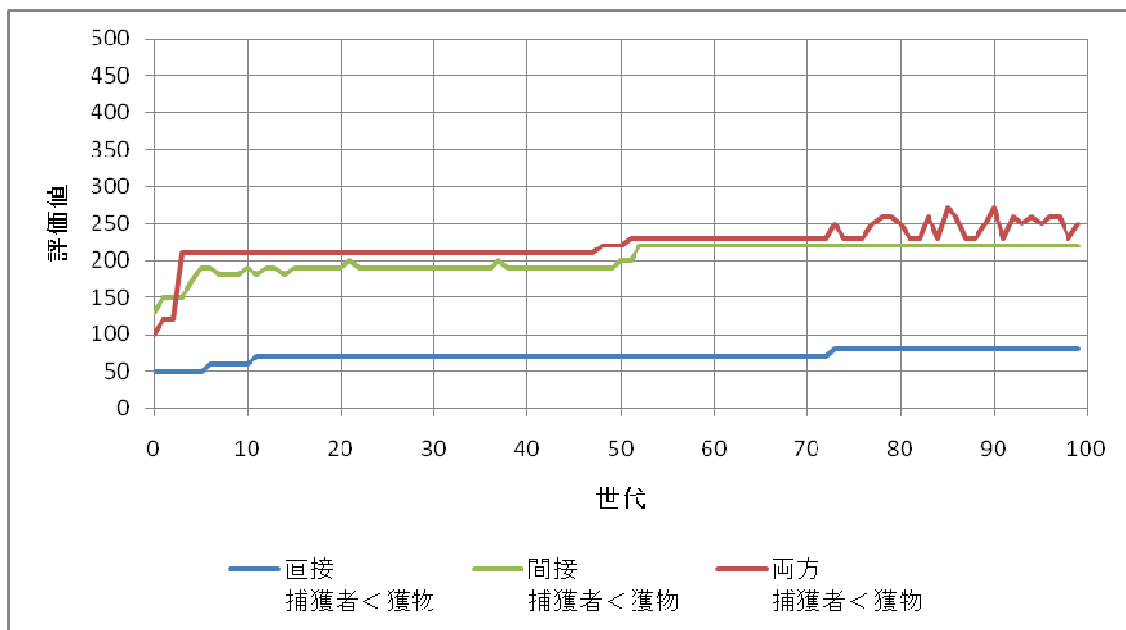


図 5.24 捕獲者の移動速度が遅い場合の直接捕獲、間接捕獲、両方の評価値

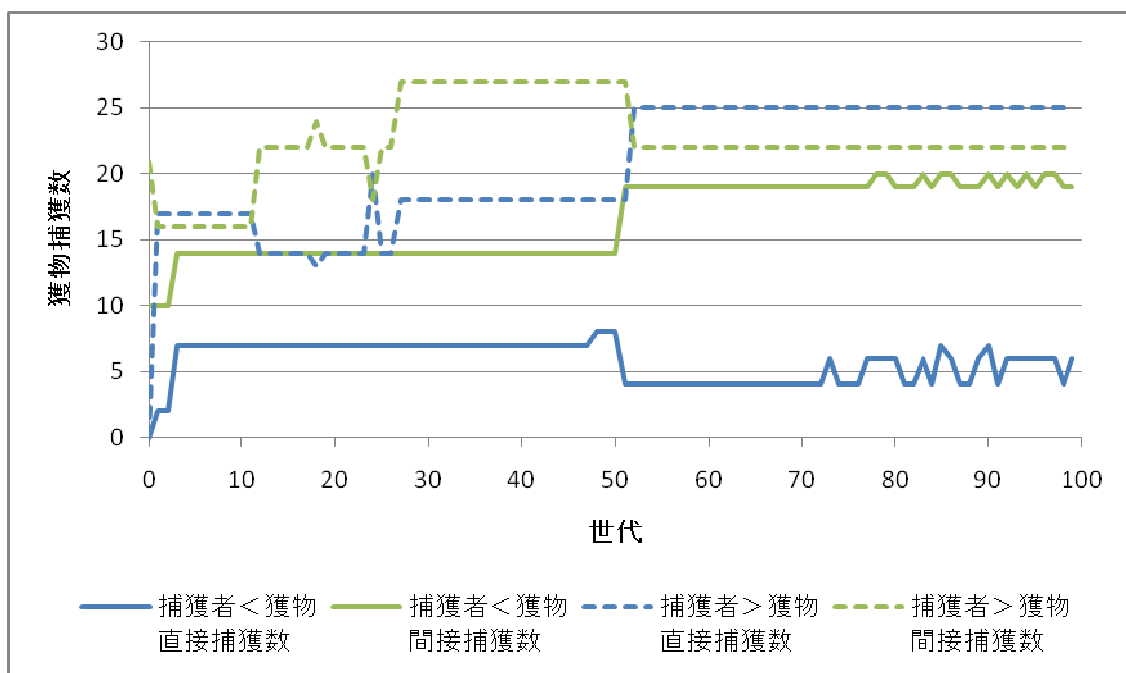


図 5.25 獲物捕獲数の推移

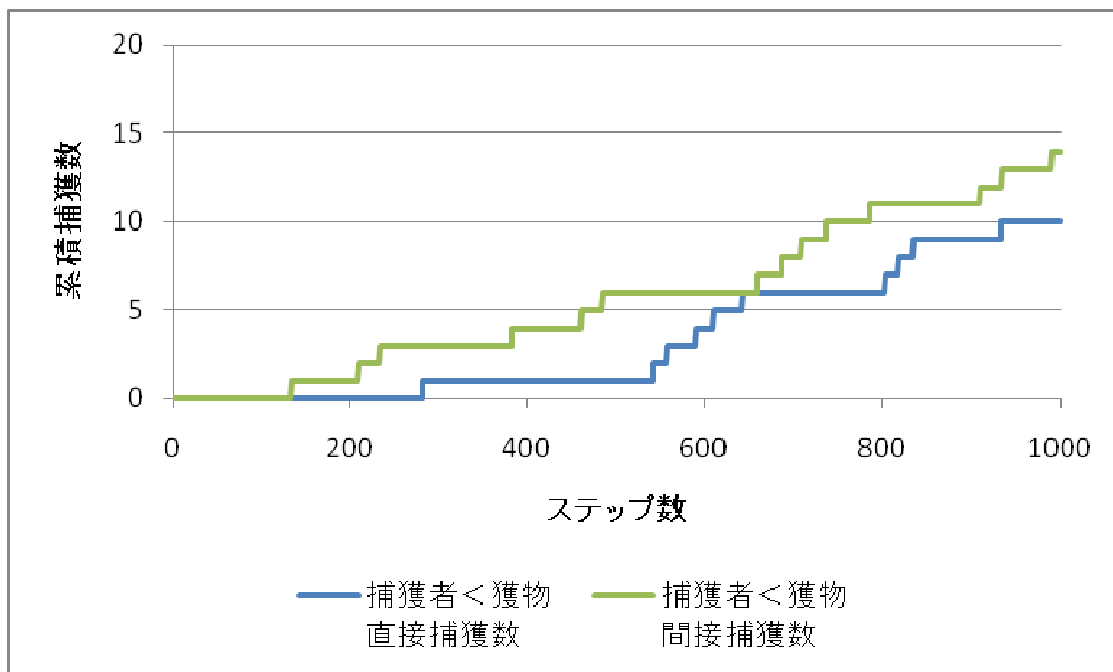


図 5.26 最終世代の試行における獲物捕獲数の推移（回避行動無し、捕獲者の移動速度が遅い場合）

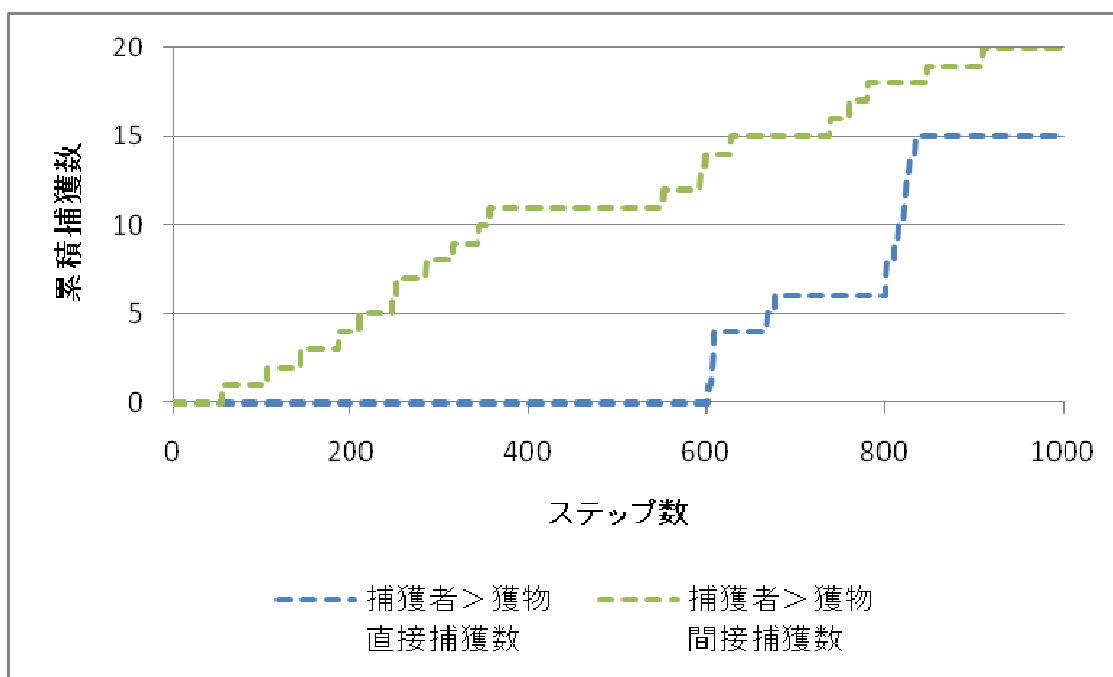


図 5.27 最終世代の試行における獲物捕獲数の推移（回避行動無し、捕獲者の移動速度が速い場合）



図 5.26、5.27 は最終世代のエリート個体における試行でのステップ毎の獲物捕獲数の推移を示したものである。図 5.26 は捕獲者の移動速度の方が獲物よりも遅い場合、図 5.27 は捕獲者の移動速度の方が獲物よりも速い場合を示している。いずれの場合も、間接捕獲数に追随して直接捕獲数が上昇している事がわかる。これは、間接捕獲によって足止めした獲物を直接捕獲している事を示しており、間接捕獲が直接捕獲の効率の上昇、ひいては捕獲全体の効率の上昇を行っている事を示している。

## 5.6.6 実験結果 2 獲物がブロック回避行動を持つ場合

### 5.6.6.1. 直接捕獲と間接捕獲を別々に行った場合の結果

図 5.28 は、直接捕獲のみを行う事ができる捕獲者を用いた場合の結果と、間接捕獲のみを行う事ができる捕獲者を用いた場合の結果を示すものである。通常線は、捕獲者の移動速度が獲物よりも遅い場合を示し、点線は捕獲者の移動速度が獲物よりも速い場合を示す。図 5.22 との違いは、獲物が障害物回避行動を持っている為、間接捕獲のコストが上昇している点である。間接行動の評価値の推移を見てみると、移動速度の違いによる評価値の違いはほとんどない事がわかるが、障害物回避行動を持たない獲物を用いた実験の時とは異なり、捕獲した獲物の数がかなり減少している事がわかる。

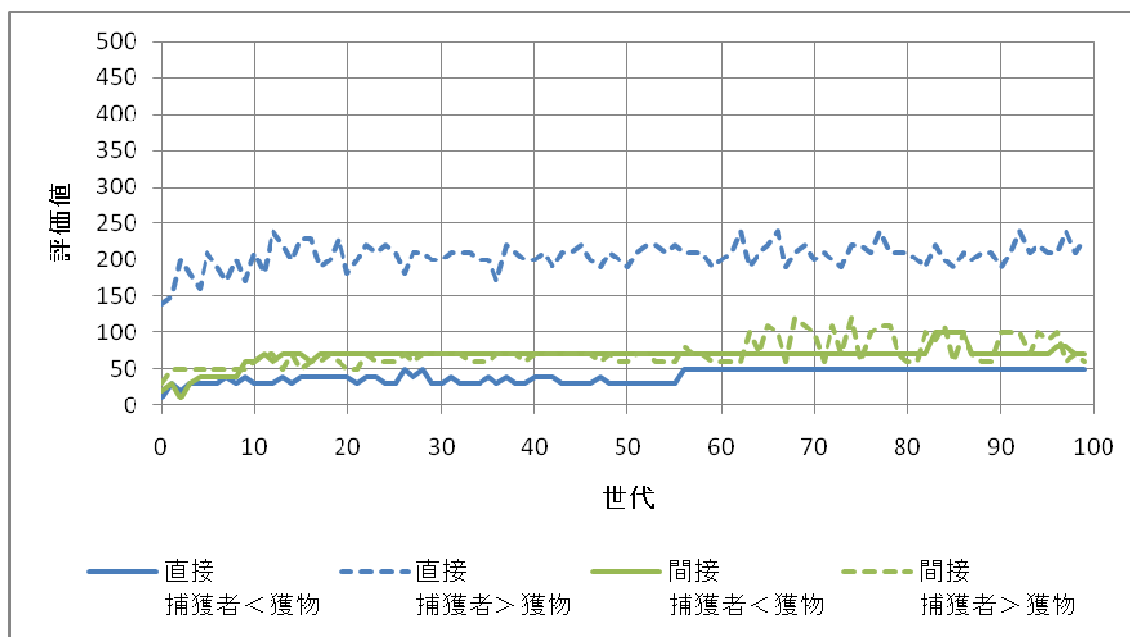


図 5.28 直接捕獲のみと間接捕獲のみでの実験結果

#### 5.6.6.2. 直接捕獲と間接捕獲のどちらも選択できる場合

図 5.29 は捕獲者の移動速度が獲物の移動速度よりも速い場合の結果、図 5.30 は捕獲者の移動速度が獲物の移動速度よりも遅い場合の結果を示している。捕獲者の移動速度が速い場合は、直接捕獲のみの場合に比べて直接捕獲と間接捕獲両方行える場合よりも評価値が低くなっている。それに対して、捕獲者の移動速度が遅い場合は、直接捕獲と間接捕獲の両方を行える場合が最も高い評価値を獲得出来ている。図 5.31 や図 5.33 に示されている捕獲獲物数の推移を見ると、捕獲者の移動速度が速い場合は、両方の行動が出来る捕獲者は直接捕獲のみを行い、間接捕獲を一切行っていない。それとは反対に、捕獲者の移動速度が獲物の移動速度よりも遅い場合は、間接捕獲を用いて直接捕獲の効率を上昇させている事が伺える。

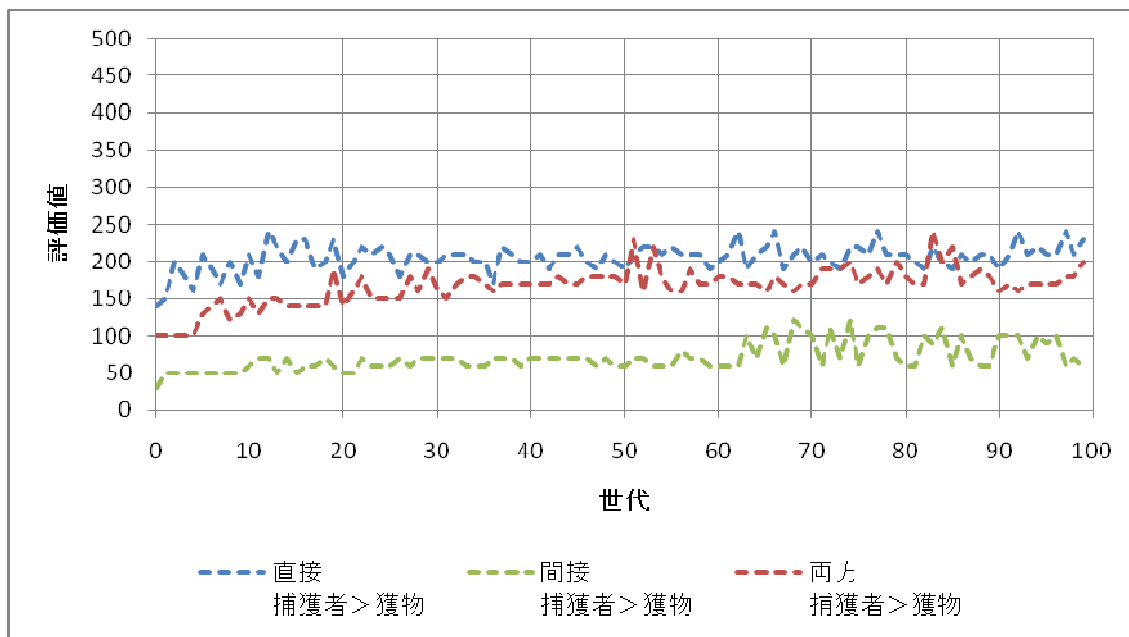


図 5.29 捕獲者の移動速度が速い場合の直接捕獲、間接捕獲、両方の評価値

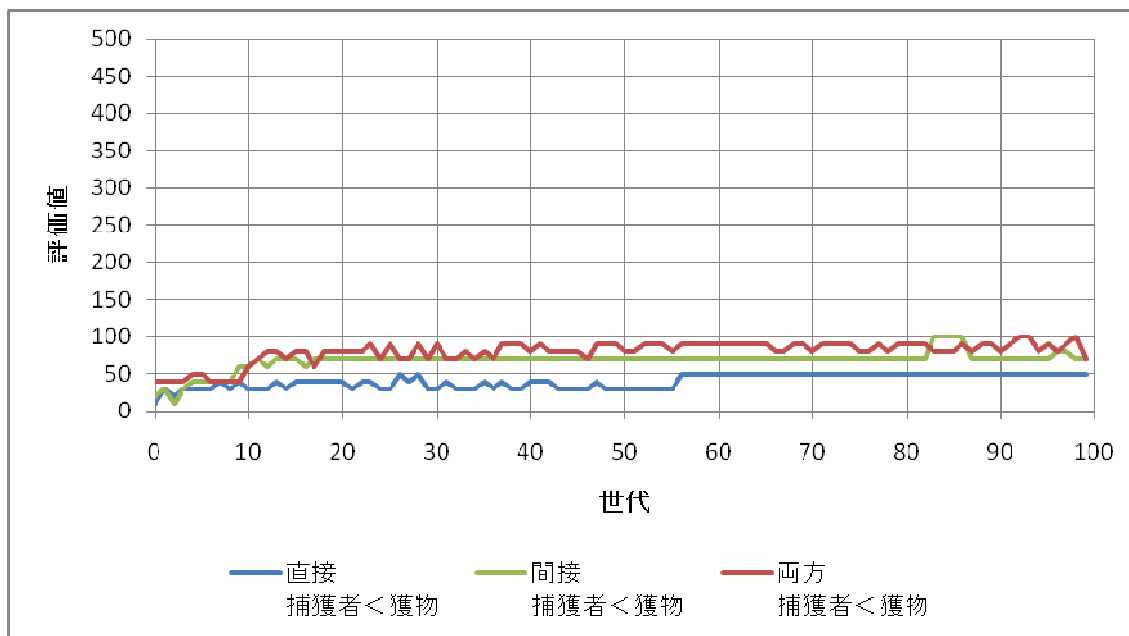


図 5.30 捕獲者の移動速度が遅い場合の直接捕獲、間接捕獲、両方の評価値

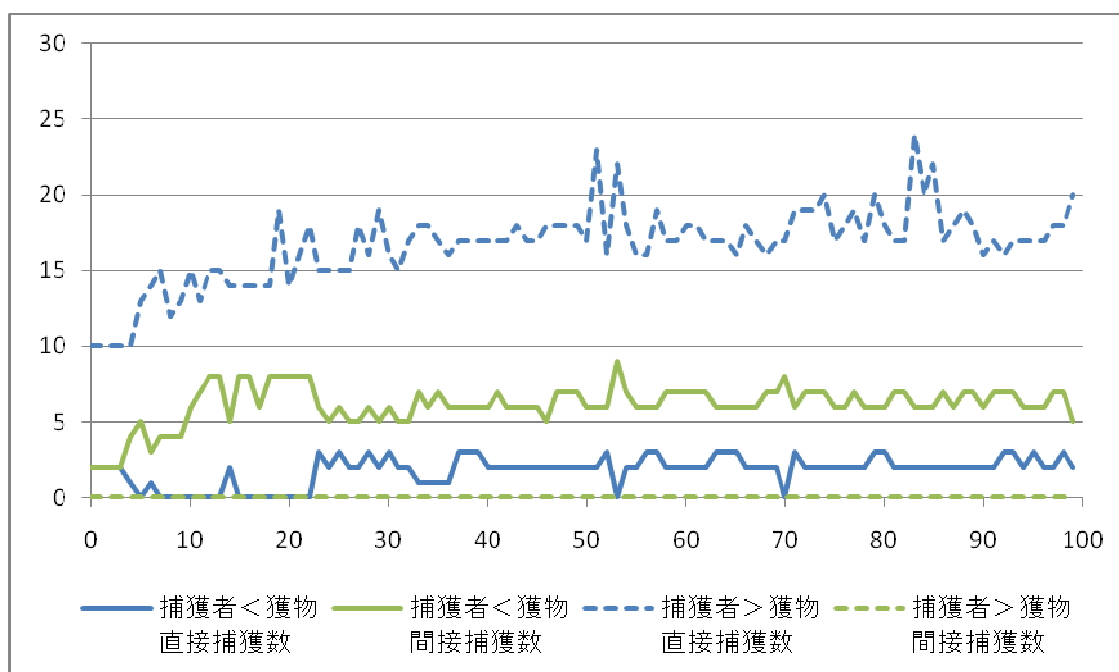


図 5.31 獲物捕獲数の推移

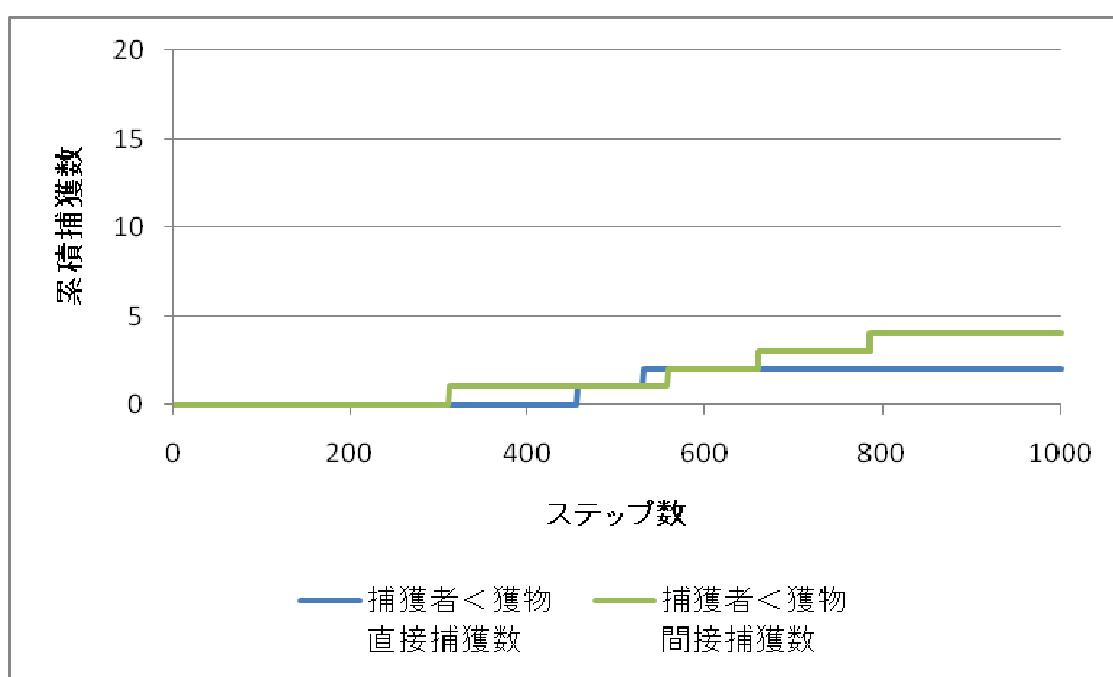


図 5.32 最終世代の試行における獲物捕獲数の推移（回避行動有り、捕獲者の移動速度が遅い場合）

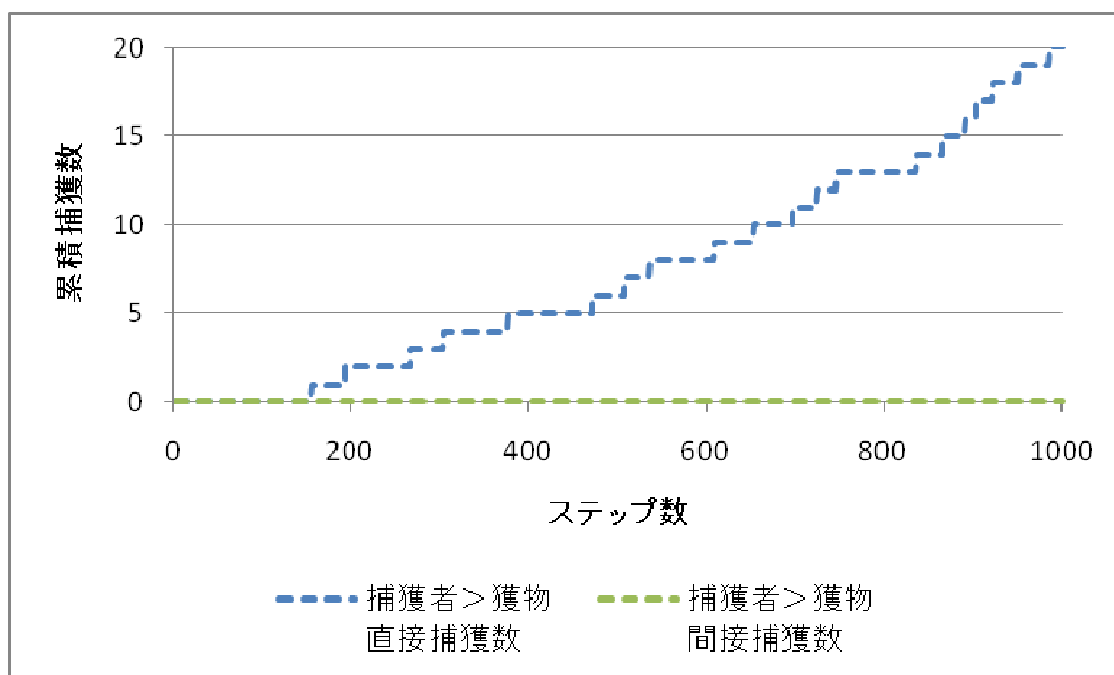


図 5.33 最終世代の試行における獲物捕獲数の推移（回避行動有り、捕獲者の移動速度が速い場合）

### 5.6.7 直接捕獲と間接捕獲の選択

表 5.2 に、表 5.1 に対して実験 2 における結果のまとめを記入したものを示す。上記の結果から、獲物と捕獲者の移動速度の関係や獲物の障害物回避行動の有無は、獲物捕獲のコストや効率に還元されると言える。即ち、それらの獲物と捕獲者の関係を表す要素は、その環境における各行動間のコストを表しており、コストが高い程その行動が獲得されにくいと考えられる。上記の結果では、直接捕獲のコスト  $C_D <$  間接捕獲のコスト  $C_I$  である時は、直接捕獲の獲得のみが行われ、直接捕獲のコスト  $C_D >$  間接捕獲のコスト  $C_I$  である時は、間接捕獲の獲得が行われている事がわかる。但し、直接捕獲のコスト  $C_D >$  間接捕獲のコスト  $C_I$  という状況では、間接捕獲の獲得が進むにつれて直接捕獲のコストが

低下する為、間接捕獲と直接捕獲を組み合わせた行動が獲得されるのだと考えられる。その証拠に、図 5.26、5.27、5.32 などですす 1 試行中における累積捕獲個体数の推移を見てみると、間接捕獲数の上昇に追従するように直接捕獲数が上昇している。これは、間接捕獲によって足止めされた獲物を直接捕獲している事を示している。

表 5.2 実験 2 の結果のまとめ

環境条件		捕獲者の獲得する行動		
獲物の回避 行動の有無	移動速度	直接行動のみ	メタ行動のみ	両方
無し	捕獲者 > 獲物	速度差に 依存する	速度差に 依存しない	直接行動とメ タ行動が交互 に進化
	捕獲者 < 獲物		捕獲量は低い	メタ行動のみ が進化
有り	捕獲者 > 獲物	速度差に 依存する	速度差に 依存しない	直接行動のみ が進化
	捕獲者 < 獲物		捕獲量は高い	メタ行動と直 接行動がわず かに進化

クモやアリジゴクなどは間接捕獲と直接捕獲を組み合わせる事で、直接捕獲のみの場合よりもより多くの獲物を捕まえる事が出来ている。これは、間接的行動によって構築した罫によって獲物を足止し、直接捕獲にかかるコストを低下させる事によって、間接捕獲と直接捕獲の組み合わせの獲得が可能であるからと考えられる。つまり、自然の例においても、直接的行動のサポートを行う間接的行動であるメタ行動を導入する事で、直接行動のさらなる改善が行われていると言える。実験 2 を通じて、それと同じ状況が 2 次元シミュレーション上でも発生しうる事が確認できた。

## 5.7 第 5 章まとめ

2 次元離散環境において、獲物捕獲用の造作物構築行動の獲得実験を行い、造作物構築行動の獲得を示す事ができた。即ち、これにより、ある程度複雑な環境下において、進化計算的手法を用いる事で、適切なメタ行動の獲得が可能であることを示す事が出来た。また、直接的行動と間接的行動の両方を行う事のできる環境において、どのような条件でどちらの行動が獲得されるのかについて、行動のコストの変化という観点から考察を行った。また、間接的行動の導入によって直接的行動の効率を上昇させうる可能性を示す事ができた。

## 6. 3次元物理シミュレーションにおける仮想生物

本章では、第5章で得られた結果をもとに、3次元物理シミュレーション上、即ち3次元連続空間上における仮想生物での構造物構築行動の獲得を試みる。実験の内容としては、2次元離散環境のシミュレーションと同じく、2種類のエージェントによる捕獲用の造作物構築行動の進化的な獲得である。Open Dynamic Engine[9, 10]を用いて3D物理シミュレーション環境の構築を行い、実験を行った。

### 6.1 シミュレーションの設定

シミュレーションの流れは、2次元環境の時と同じものを採用している。即ち、図5.5に示す通り、遺伝的アルゴリズムシミュレーションのフローを基本として、意思決定機構に人工ニューラルネットワークを用い、その荷重を遺伝子によって決定している。各遺伝子について試行を行い、最も評価の高かった遺伝子をもとに次世代の生成を行う。今回のシミュレーションでは、その試行部分を3Dシミュレーション環境で構築している。

#### 6.1.1 シミュレーション環境

図6.1に3D物理シミュレーションの環境の概要を示す。環境の中心には、捕獲エリアがあり、獲物エージェントが捕獲エリアに侵入した場合、その獲物エージェントは捕獲されたとみなされ、以降行動不能となる。これは、クモの巣やアリジゴクの巣のように一定の場所で狩りを行う動物を想定した設定である。



環境の一端には、獲物エージェントの環境への流入口であるスタートエリアがある。その反対側には、獲物エージェントの環境からの流出口である脱出エリアがある。獲物エージェントの目的は、環境から脱出することである。捕獲者エージェントの目的は、獲物エージェントを捕獲することである。但し、その捕獲方法として、ブロックを用いて獲物エージェントをスタック状態にするか、捕獲エリアへと誘導するか、いずれかの方法を実現する必要がある。

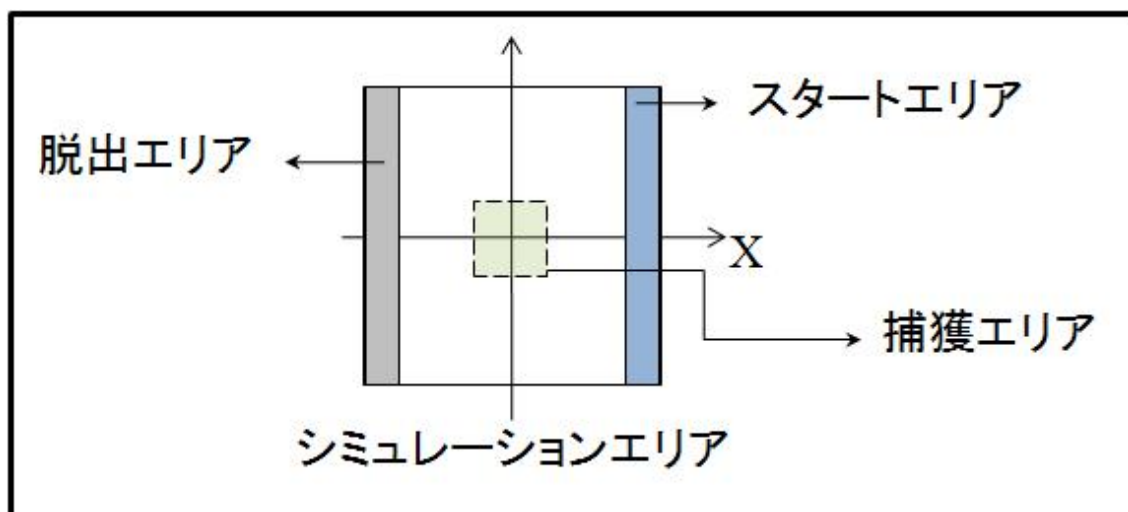


図 6.1 3D 物理シミュレーション環境の概要

### 6.1.2 獲物エージェント

3D シミュレーション環境における獲物エージェントは、単純反射エージェントの一種であり、簡単な障害物回避ルールを持つ。図 6.2 に獲物エージェントの外観を示す。獲物エージェントは 2 つの車輪、シャーシ、頭部を持ち、センサとしてビジョンセンサを備えている。獲物エージェントはビジョンセンサによってブロックの位置を認識することが出来、ブロックの回避を行う事が出来る。図 6.3 にブロック回避の様子を示す。眼の前にブロックがある場合、ブロックの

無い方向へ旋回しつつ移動を行う事で、ブロックの回避を行う。ブロックの位置によっては、獲物エージェントはブロックを回避することが出来ず、スタック状態に陥ってしまう。

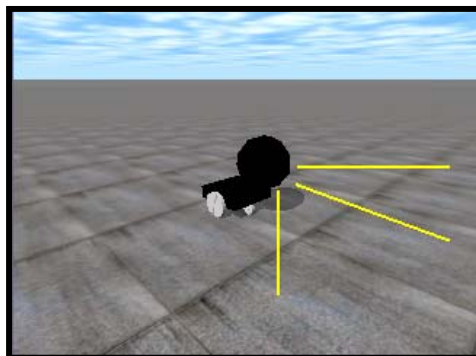


図 6.2 獲物エージェントの外観

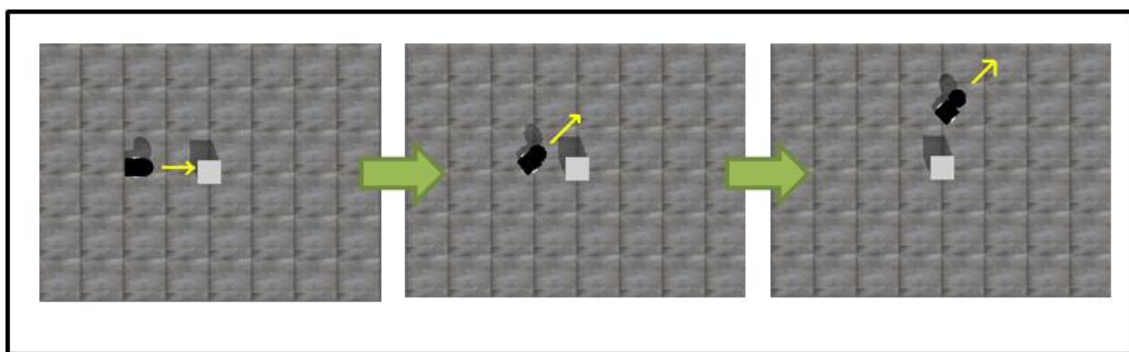


図 6.3 障害物回避の様子

### 6.1.3 捕獲者エージェント

図 6.4 に捕獲者エージェントの身体構造を示す。捕獲者エージェントの身体は頭部、首、シャーシ、二つの車輪から構成されている。センサとして最大 8 個の近接センサを持つ。近接センサは一定の長さを持ち、センサと接触したブロ

ックとセンサの根元までの距離を値として返す。ブロック未検出時には、センサ最大値にマイナスを掛けたものが返される。内部状態としては、2次元シミュレーション環境時と同じく、現在の向き、現在位置、設置可能ブロック数を持つ。行動としては、直進、左旋回、右旋回、ブロックの設置の4種類の行動を持つ。これらの行動は、2次元シミュレーション環境時と同じ内容になっている。

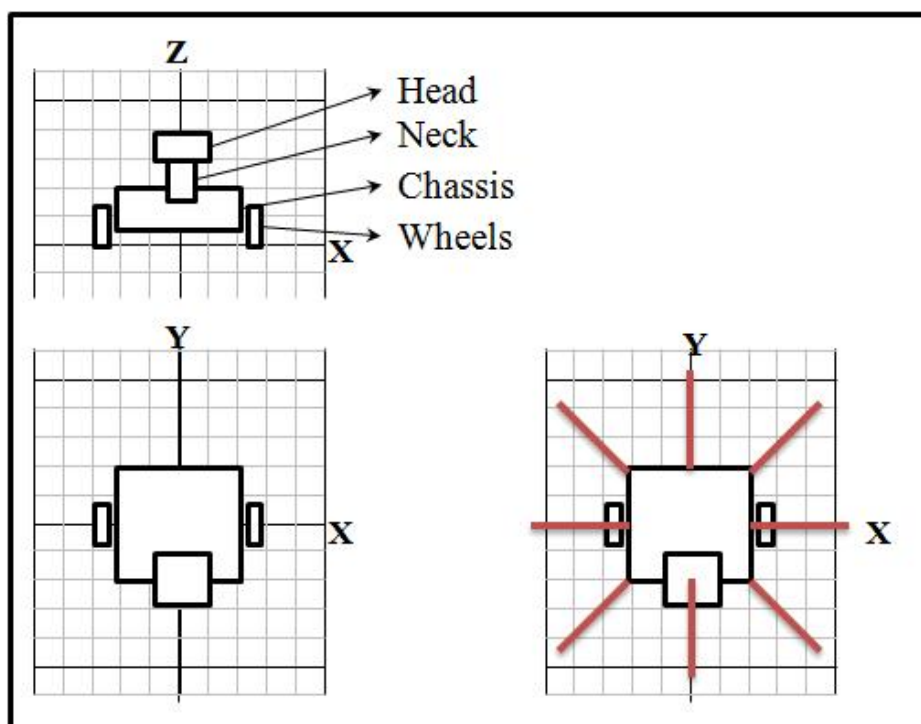


図 6.4 捕獲者エージェントの身体構造

#### 6.1.4 意思決定機構

捕獲者エージェントは意思決定機構として2層ニューラルネットワークを採用している。基本的には2次元シミュレーション環境時と同じだが、環境情報に当たる部分が近接センサからの値へと変更になっている。

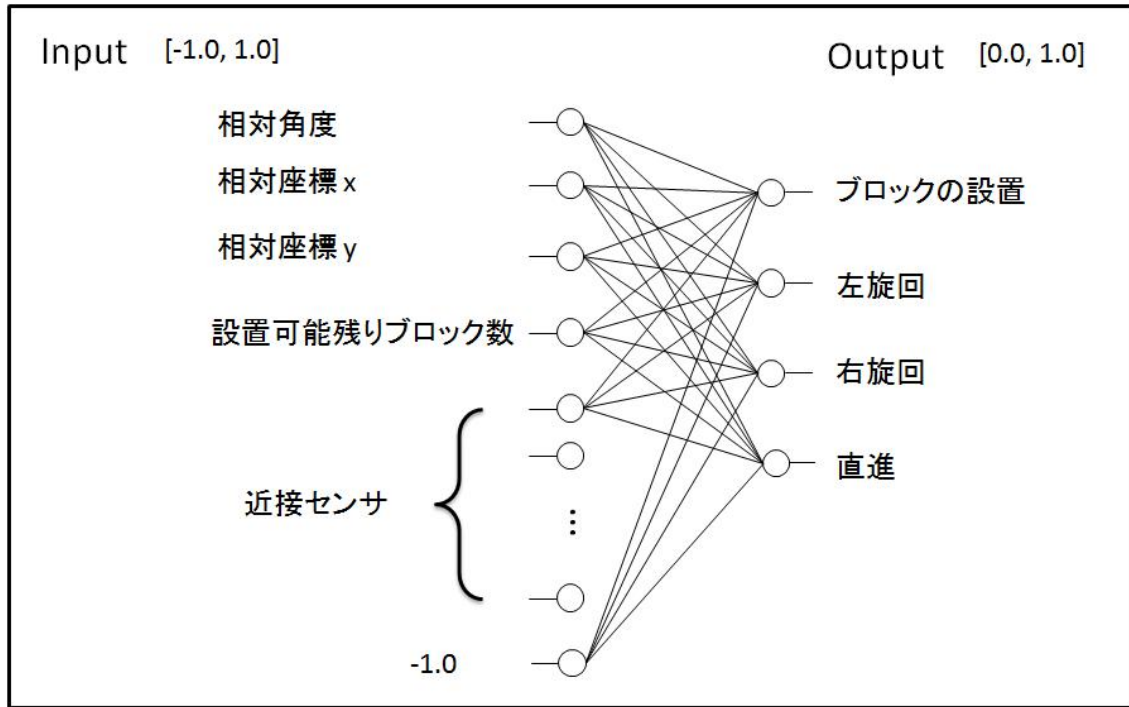


図 6.5 3D 物理シミュレーション環境における捕獲者エージェントの意思決定機構

#### 6.1.5 評価関数

構造物の評価として、下記の評価関数を用いた。

$$Fitness = \alpha T + \sum_j^{S-T} \beta \frac{1}{D_j} \quad (6.1)$$

$\alpha$ 、 $\beta$  は定数、 $T$  は捕獲することのできた獲物エージェントの数、 $S$  は獲物エージェントの全体の数である。 $D_j$  は  $j$  体目の獲物エージェントと中心との距離である。この評価関数では、第 1 項では捕まえた獲物の数による評価を、第 2 項では、捕獲できなかった獲物エージェントをどれだけ中心部に近い位置に留める事が出来たかを評価している。図 6.6 は評価の例を示している。この場合、2 体を捕獲エリアに誘導することが出来たので、 $T=2$  で第 1 項が計算され、残りの 1 体と中心部までの距離を用いて第 2 項が計算される。

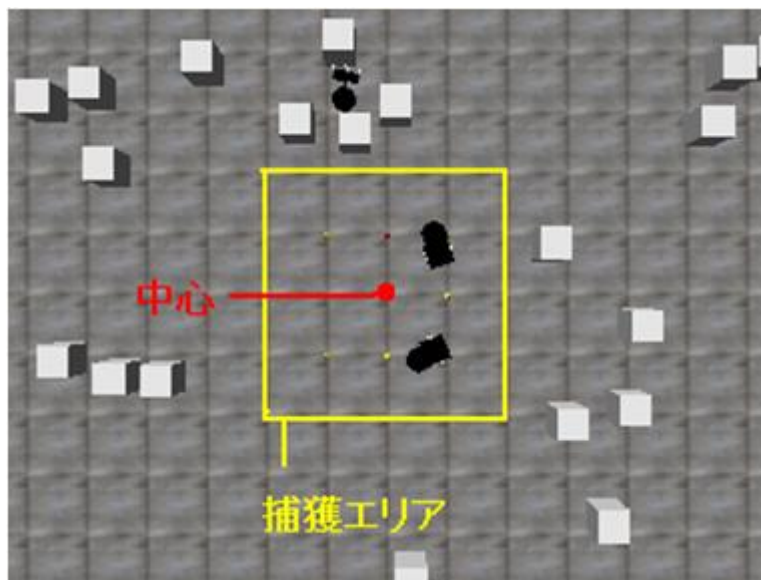


図 6.6 構造物の評価

## 6.2 実験 1 構造物構築行動の進化的な獲得

3D 物理シミュレーション環境における構造物構築行動の進化的な獲得を目的として実験を行った。また、環境のセンシング方法によって異なる行動が獲得されるかどうか検討を行う為に、近接センサの数を変えて実験を行った。近接センサの数は、正面のみの 1 つの場合、正面 3 つの場合、周囲 8 個の場合の 3 種類である。

### 6.2.1 実験パラメータ

#### 環境

- ・ 環境の広さ                    10.0[m]×10.0[m]
- ・ 獲物エージェントの数        16 体
- ・ 最大ブロック数            5 個

- ・ブロックの寸法       $0.5[m] \times 2.0[m] \times 0.5[m]$

#### 遺伝的アルゴリズム関係

- ・突然変異率    0.10
- ・交叉率                      0.70
- ・保存エリート              1 体
- ・1 世代の個体数          30 体

#### 6.2.2 実験結果

図 6.7 に各センサ数におけるエリート個体の評価値の遷移を示す。センサ数が 3 の場合が最も高い評価値を獲得しているが、他の場合と比べてあまり大きな差がない事がわかる。センサ数 8 の場合が最も評価が低いが、世代数が足りず進化が不十分である可能性がある。評価値としては、3～6 体の獲物エージェントを捕獲出来ている事が伺える。従って、獲得された構造物は 16 体中の半分以下程度の獲物エージェントしか捕獲することが出来ておらず、ある程度しか環境へ適応できた巣を作る事が出来ていない事がわかる。

図 6.8 に、各センサ数において、獲得された構造物の例を示す。各場合において、獲得された行動は基本的には同一種類のものであった。その場で回転しながら、ブロックを設置していくという行動である。センサ数 1 の場合は、ブロックを設置して回転し、センサにブロックの反応がなくなると、またブロックを置くという行動を繰り返している。センサ数 3 の場合も同様であった。センサ数 8 の場合は、1 つブロックを置いたあと、しばらく回転を続け、その後再度ブロックを置き始めた。これは、ブロックを検知するセンサが複数あり、それ

らが離れた位置関係であった為だと思われる。しかし、その結果として、片方のセンサがブロックを検知して回転する方向と、もう片方のセンサがブロックを検知して回転する方向が逆であった為、その場で向きの変換を繰り返すスタック状態に陥っており、結果としてブロックを 3 個までしか配置することが出来なくなっていた。センサ数 8 の場合、用いるセンサ数が増えた為、行動間の競合が発生しやすくなっている事が推測される。その為に、他の場合にくらべて、適切な行動を獲得するまでに必要な世代数が増え、結果として今回の設定では、環境に適応した行動が獲得できなかったと考えられる。

いずれの場合にしても、獲得された行動は環境に対してある程度適応できてはいたが、不十分であると言える。その為、本研究では、自然の動物の構造物構築行動、特にクモの造網行動に注目し、その特徴の一つである中間評価について着目した。

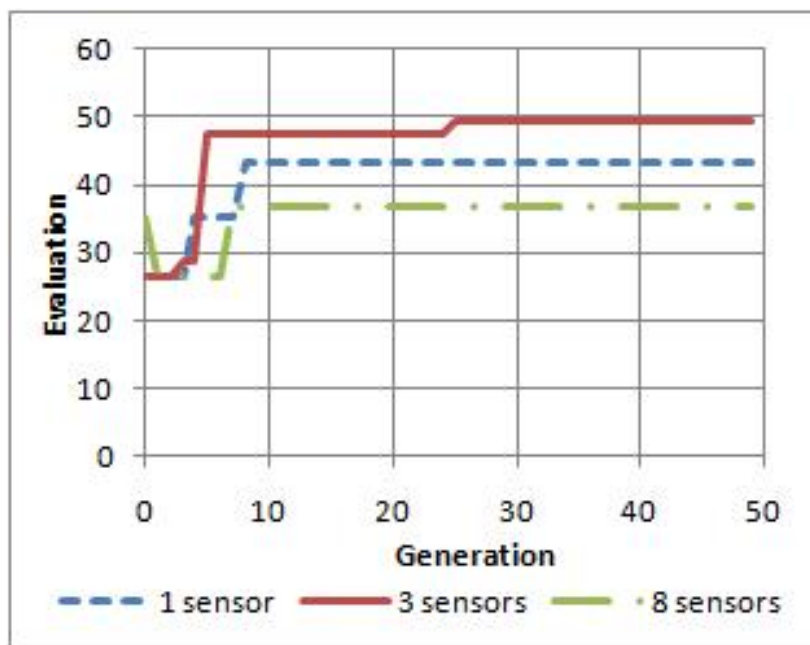


図 6.7 異なる近接センサ数におけるエリート個体の評価値の遷移



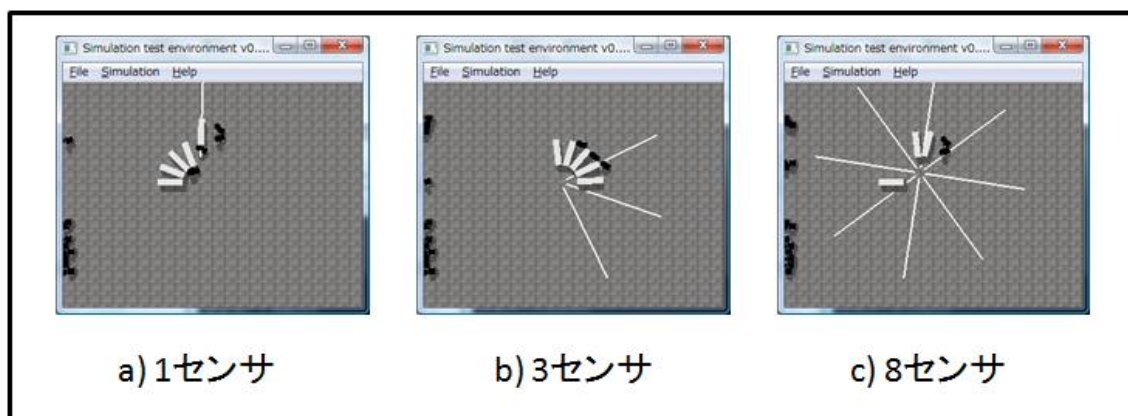


図 6.8 獲得された構造物の例

### 6.3 実験 2 中間評価の影響

実際のクモの造網行動を鑑みてみると、クモは造網行動の最中に巣の張力を何度も確かめ、意思決定の重要な情報としていと考えられる。例えば、一本の放射糸を張る際に、クモは必ずこしきに戻り、巣の張力を確かめる。張力が不十分であれば更に追加の放射糸を張り、十分であれば次のプロセスである足場糸掛けへと移行する。また、他のプロセスにおいても、張力が極端に変化した場合巣の異常を察知し、通常とは異なる行動をとる場合がある。このように、巣の状態を張力によって検知し、意思決定の重要な情報として扱っている事がわかる。以上の事柄から、実際の生物における構造物構築行動の特徴について以下の点が挙げられる。

- 1) 自身の構造物の特徴を評価し、その結果を意思決定に用いている点
- 2) 巣の物理的性質と張力を検知できる身体的能力を用いて、巣全体の特徴を張力という情報量に変換している。即ち、環境と身体との物理的な相互作用を通じてグローバルな情報をローカルな情報に変換し、意思決定に用いている



今回の実験では、(1)を考慮し、意思決定機構への入力情報に構造途中の中間における構造物の評価を明示的に用いる事で、獲得される行動に対する影響を確認する。(2)に対しては、今回の実験では明示的に含める事はしない。

### 6.3.1 中間評価

今回の実験では、捕獲者エージェントの周囲のブロック数を中間評価として採用した。式は以下の通りである。

$$sEval = \begin{cases} N_b & (if\ N_b < MAX_B / 2) \\ MAX_{Eval} & (otherwise) \end{cases} \quad (6.2)$$

但し、 $N_b$ は捕獲者エージェントが検知したブロックの数、 $MAX_B$ は設置可能ブロックの最大数、 $MAX_{Eval}$ は定数である。即ち、 $N_b$ が最大値の半分以下である場合は、 $N_b$ が評価となり、最大値の半分以上である場合は一定の評価値が与えられる。

### 6.3.2 意思決定機構への変更

図 6.9 に中間評価を入力の一部として明示的に追加した意思決定機構について示す。前の実験で用いられた 2 層 ANN に対して、入力ノードを 1 つ追加し、その入力として上記の中間評価関数によって計算された値を用いている。

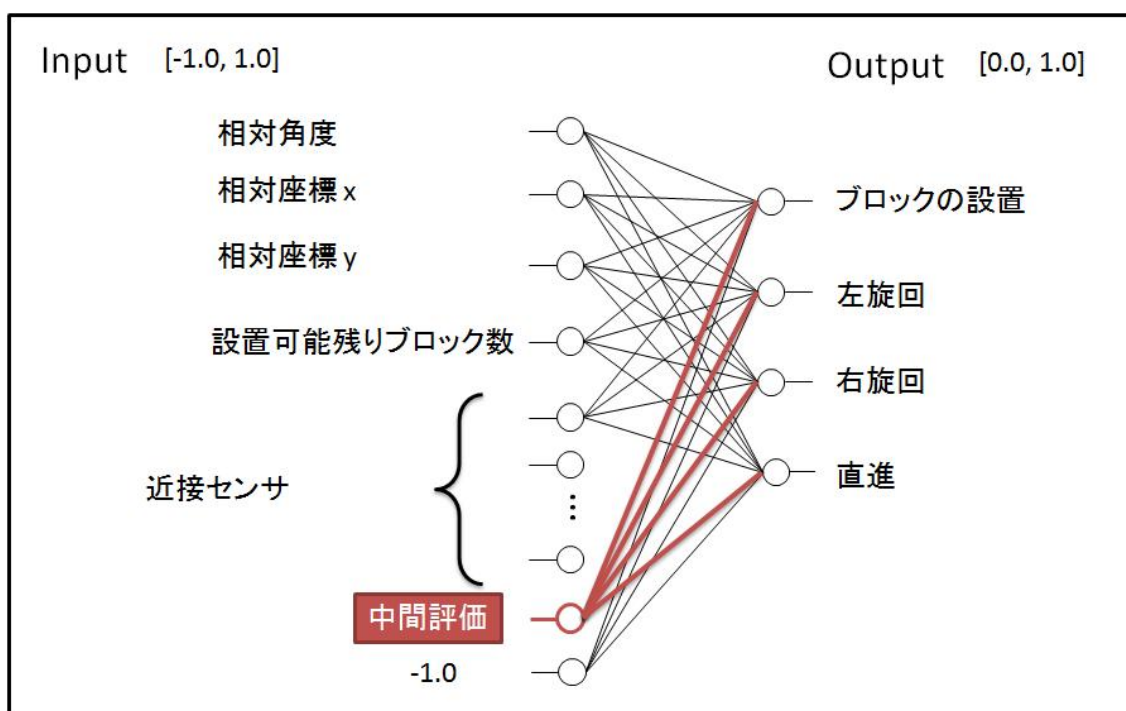


図 6.9 中間評価を加えた意思決定機構

### 6.3.3 実験パラメータ

#### 環境

- ・ 環境の広さ                     $10.0[\text{m}] \times 10.0[\text{m}]$
- ・ 獲物エージェントの数        16 体
- ・ 最大ブロック数            10 個
- ・ ブロックの寸法             $0.5[\text{m}] \times 2.0[\text{m}] \times 0.5[\text{m}]$

### 6.3.4 実験結果

図 6.10～6.12 に各センサ数におけるエリート個体の評価値の遷移を示す。センサ数 1 の場合は 50 世代で評価値が一致しているが、センサ数 3 の場合は中間

評価なしの方がわずかに高く、センサ数 8 の場合は、中間評価ありの方がわずかに高くなっている。これらの結果から考えるに、進化の速度については特に差異はないと考えられる。2 層ニューラルネットワークの構造から考えても、差は 1 つの入力ユニットとその荷重 4 本のみなので、妥当な結果だと考えられる。

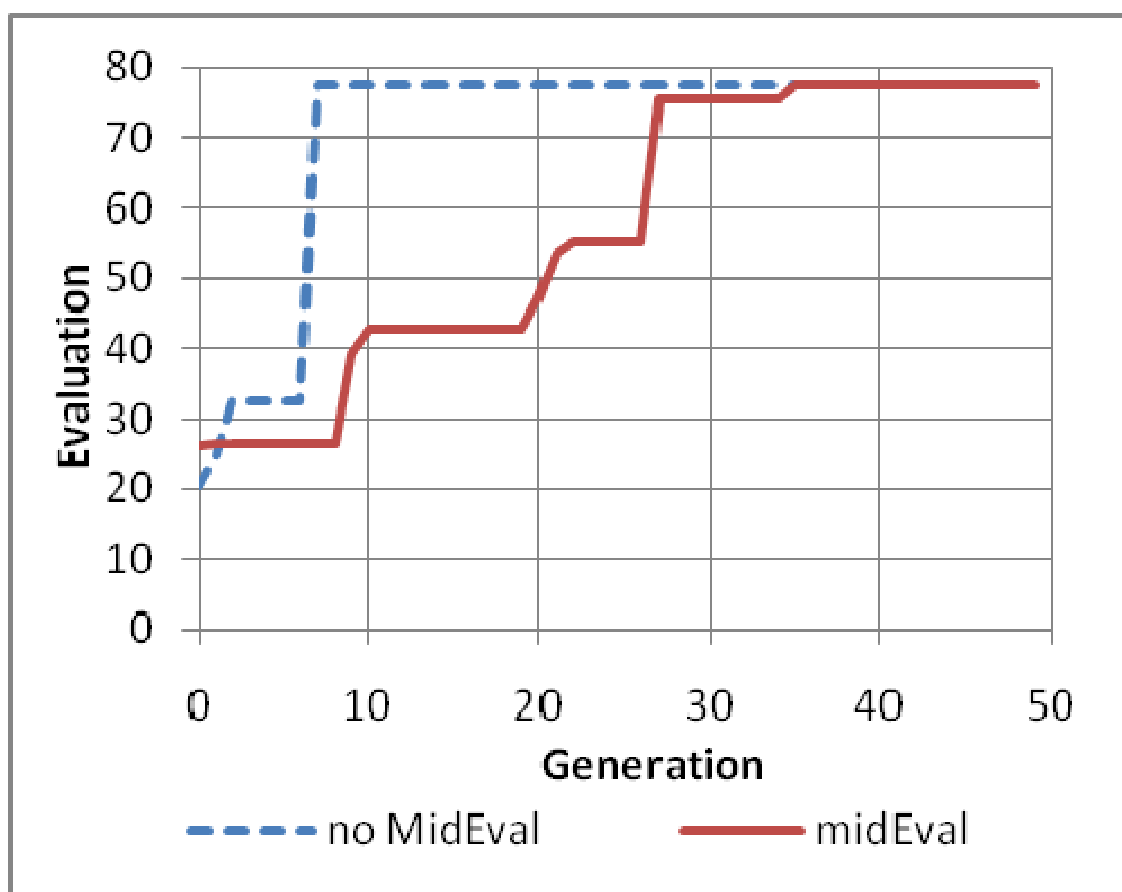


図 6.10 センサ数 1 の場合のエリート個体の評価値の遷移

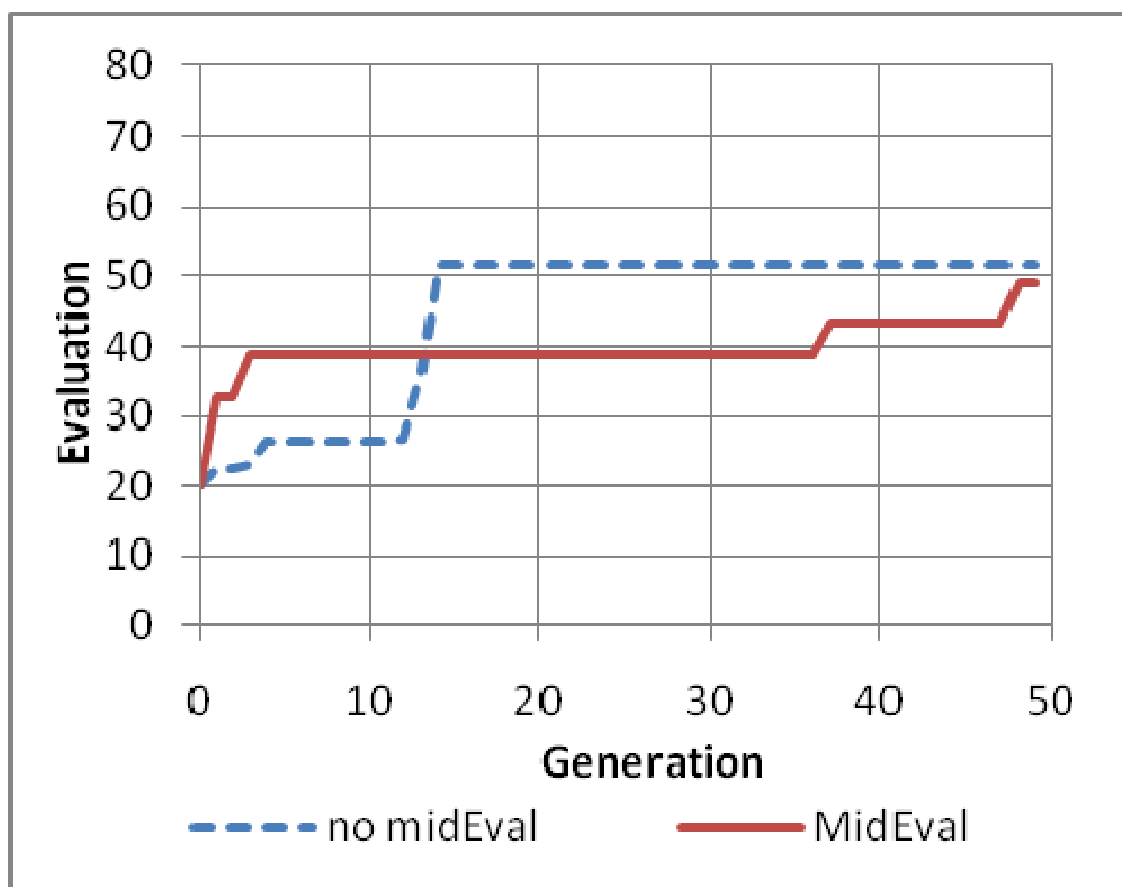


図 6.11 センサ数 3 の場合のエリート個体の評価値の遷移

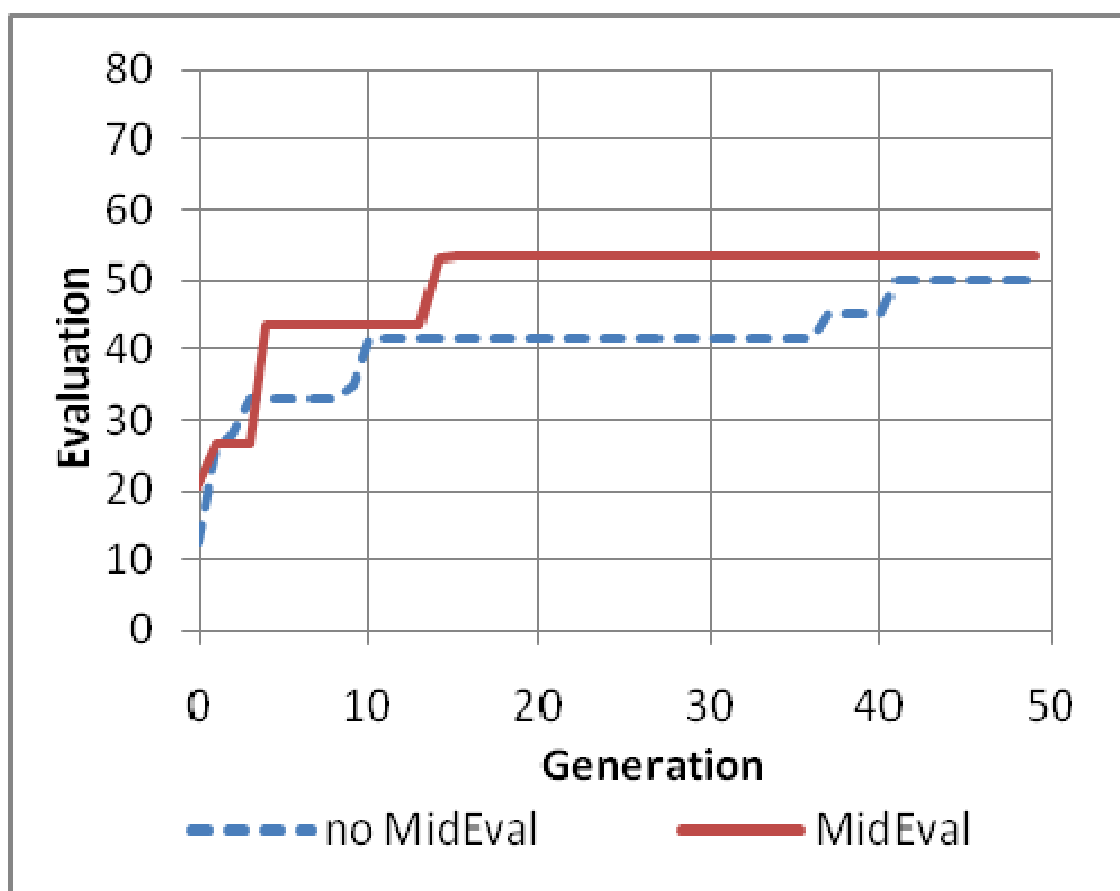


図 6.12 センサ数 8 の場合のエリート個体の評価値の遷移

図 6.13～6.15 に各センサ数において獲得された構造物構築行動の様子を示す。図 6.13 の左側は中間評価なしの場合の構造物構築行動を、右側は中間評価ありの場合の構造物構築行動を示している。中間評価なしの場合、捕獲者エージェントは、ブロックを置きながらその場で回転を行う。そして、五つ目のブロックを置いた後、設置した直後のブロックに沿いながら移動を開始し、ブロックの反応が消えると再度、新しいブロックを置くという行動を行う。中間評価ありの場合、獲物エージェントは、ブロックを置きながら前進を行い、ブロックの積み上げを行っていく。

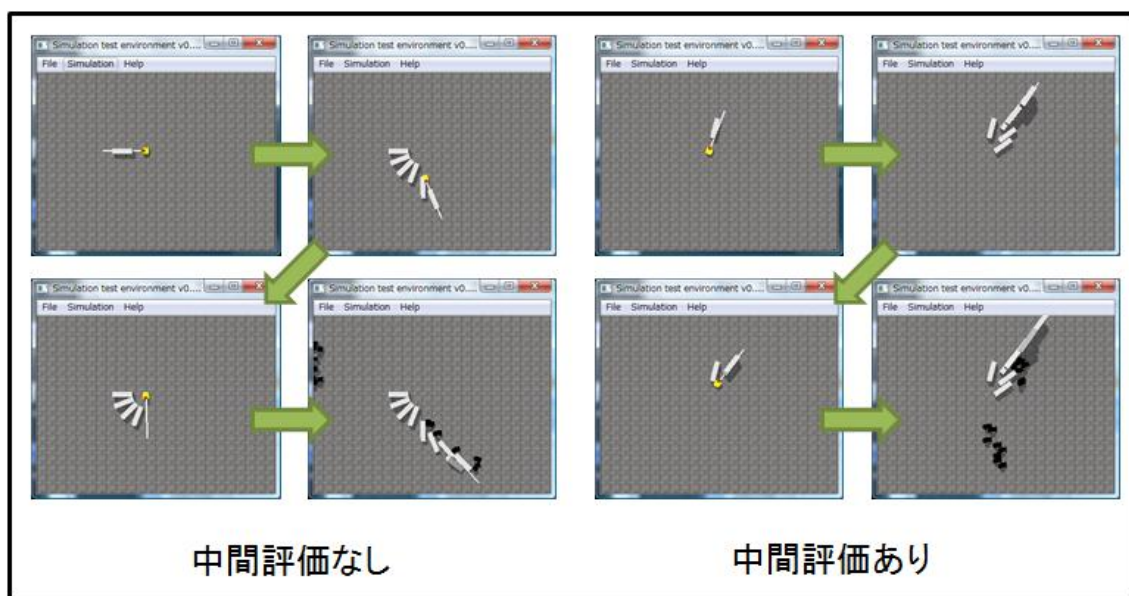


図 6.13 センサ数 1 の場合における構造物構築行動の様子

図 6.14 の左側はセンサ数 3 の場合の中間評価のない捕獲者エージェントの構造物構築行動を、左側は中間評価のある捕獲者エージェントの構造物構築行動を示している。中間評価を用いていない場合、獲得された行動はセンサ数 1 の場合の中間評価を用いていない捕獲者エージェントと類似している。中間評価を用いている捕獲者エージェントの場合、ブロックを設置しながら回転する行動を行った後、少し前進して、もう一度回転しながら設置している。

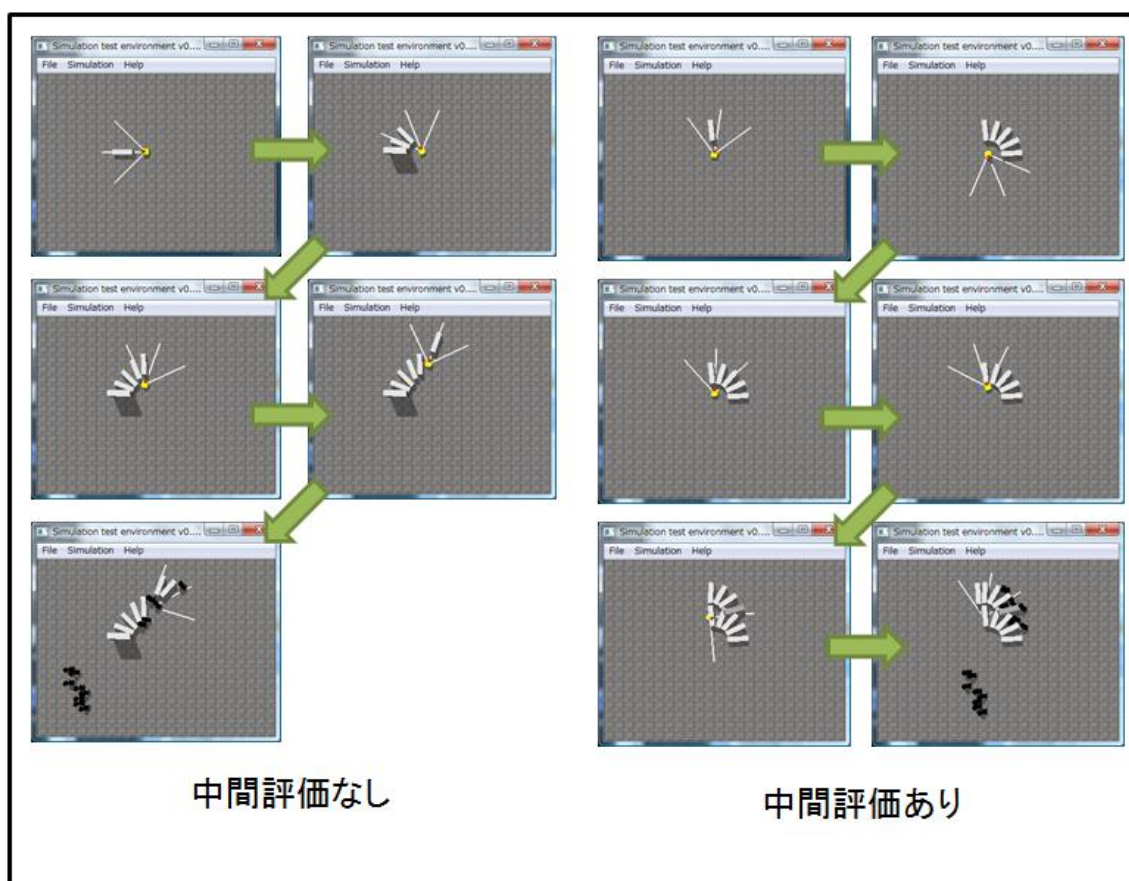


図 6.14 センサ数 3 の場合における構造物構築行動の様子

図 6.15 の左側は、中間評価を用いない場合の捕獲者エージェントの構造物構築行動を、右側は中間評価を用いた場合の捕獲者エージェントの構造物構築行動を示している。中間評価を用いていない場合の捕獲者エージェントは、その場で回転しながらブロックを設置する行動を二回繰り返している。一度目の回転で置いたブロックの間にブロックを置き、隙間を詰めるような行動を獲得している。一方、中間評価を用いている捕獲者エージェントは、回転しながらブロックを二個ずつ離れた間隔で設置していき、6 個目のブロックを置いたあと、前方へと移動し、一定の距離を進んだ後、再度ブロックを置き始めている。これは、周囲に 6 個のブロックを設置した事によって、構造物の構造評価  $sEval$  の値が行動の切り替えを誘発した事が、このような振る舞いが発生した原因の一つであると考えられる。



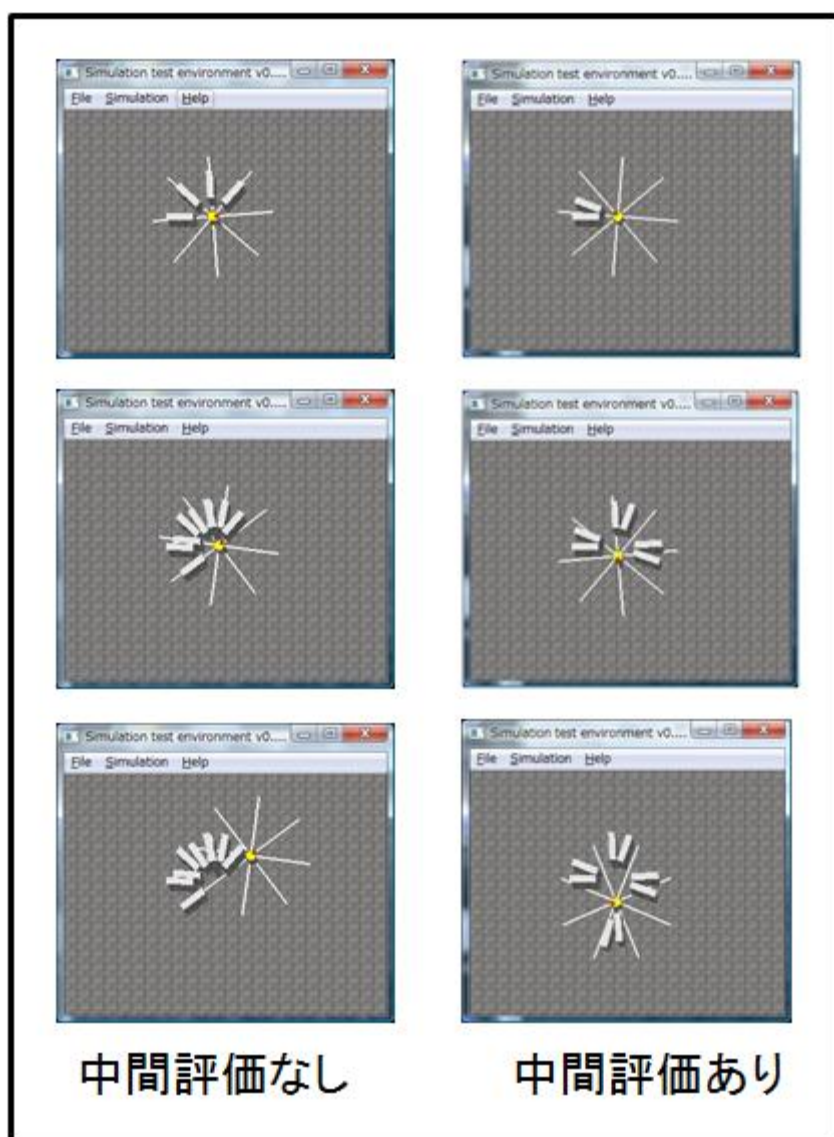


図 6.15 センサ数 8 の場合における構造物構築行動の様子

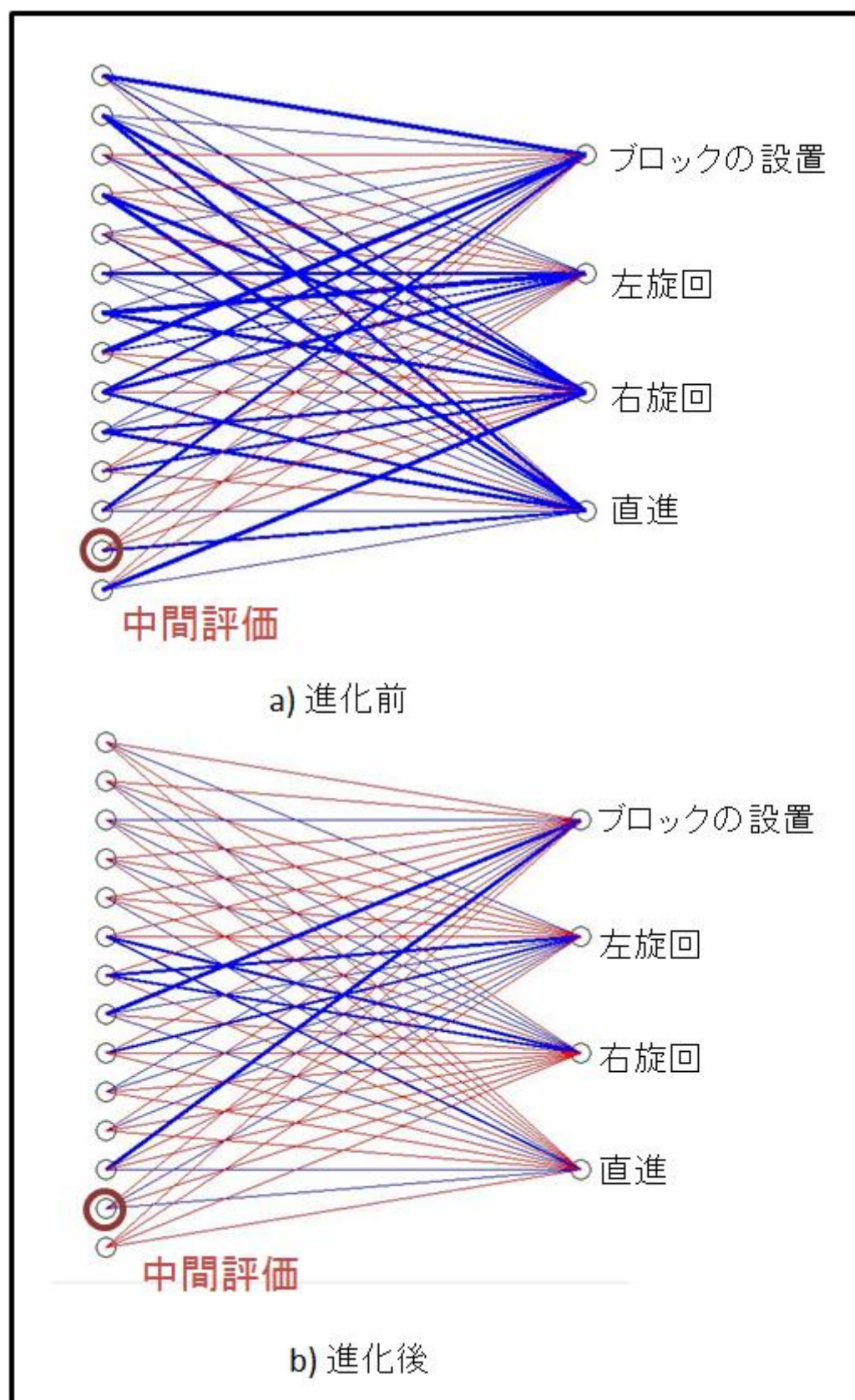


図 6.16 センサ数 8 の場合における進化前後の ANN

図 6.16 はセンサ数 8 の場合における進化前と後の人工ニューラルネットワークの様子を示している。赤い丸で囲まれた入力ユニットは、中間評価の入力に対応している。青いコネクションは正の値の荷重を示し、赤いコネクションは負の値の荷重を示している。コネクションの太さは値の大きさを示しており、絶対値が大きいほど線が太くなる。進化前の ANN は、大きな値を持つ正の荷重が多い。その一方、進化後の ANN は、大きな値を持つ荷重自体が少なくなっている事がわかる。

ここで、中間評価に対応する入力ユニットの荷重に注目してみる。進化後の ANN では、ブロックの設置、右旋回に対して負の荷重が繋がり、左旋回、直進に対して正の負荷が繋がっている。試行開始当初は、捕獲者エージェントの周りにブロックはなく、sEval の値は 0 になっている。即ち、行動の選択に関して中間評価の値が全く考慮されていない状態である。しかし、捕獲者エージェントが自身の周囲にブロックを多く配置すると、sEval の値は大きくなり、ブロックの設置、右旋回の行動は抑制され、左旋回と直進の行動が刺激される。これによって、左旋回と直進の行動が選択されやすくなると考えられる。以上のようにより、中間評価の情報を用いて重要な意思決定を行うような振る舞いが獲得されている事がわかる。

#### 6.4 第 6 章のまとめ

3D 物理シミュレーション環境において、構造物構築行動を進化的に獲得する仮想生物の枠組みについて実験を行い、自身の行動や環境に合わせた構築行動の獲得が可能であることを示唆する事が出来た。また、中間評価を入力情報の 1 つとして扱う事で、実際の生物の持つ構造物構築行動の特徴を再現することができる可能性を示唆する事ができた。即ち、環境に合わせた適切なメタ行動の獲得の可能性を示唆する事ができた。

## 7. 結論

本章では、研究としての結論を述べ、将来的な展望として、提案手法の応用として考えられる例を述べる。

### 7.1 結論

本研究では、ある目的を直接的に達成しようとする行動を直接的行動、直接的行動をサポートし、その効率性や実現可能な事柄の改善が可能である間接的な行動をメタ行動と呼び、焦点を当て、メタ行動の導入による問題解決アプローチ提案の為の基礎研究を行ってきた。メタ行動の性質について議論を行うとともに、メタ行動の導入の枠組みの有効性の検証を行う事を目的として、共有地の悲劇と造作物構築行動の獲得という 2 種類の問題に対して適用を行い、段階的に研究を進めてきた。

共有地の悲劇は、確定的なルールによってゲーム全体の挙動が予測可能であるという特徴がある。また、エージェントの行動に対してコストが考慮されず、実行によって即座に利得が獲得されるという環境であり、行動のプロセスではなく、動作が評価対象となっている。エージェントの行動は、直接行動かメタ行動かのいずれかを選択するという形式になっている。それに対して、捕獲用の造作物構築行動の獲得では、確定的なルールによってシミュレーションの挙動が決定されておらず、ランダム要素を含む為、環境の挙動の予測は難しい。また、複数の動作を組み合わせて直接捕獲や間接捕獲の一連のプロセスを獲得する必要があり、獲得したプロセスが造作物の評価を通じて間接的に評価されていると言える。行動についても、直接行動かメタ行動の選択ではなく、基本的な行動の組み合わせを用いて直接行動やメタ行動を獲得する。

以上の問題へのメタ行動の導入を通じて、その有効性を段階的に確認する事が出来た。結果として、メタ行動の進化的な獲得について、遺伝的アルゴリズムや NeuroEvolution 手法による枠組みを示す事ができた。また、直接行動だけの環境よりもメタ行動を導入した環境の方が良いパフォーマンスを得る場合がある事を示す事ができた。そして、造作物構築途中での構造的な評価を行い、意思決定の入力とする事で、行動パターンの切り替えを含むプロセスの獲得について示す事が出来た。

## **7.2 メタ行動の設計論に向けて**

本研究では、直接的な行動のサポートを行う間接行動をメタ行動と呼び、研究を進めてきた。これは、直接行動と間接行動との関係性において階層性が認められる事を考慮してのことである。しかし、一般的な定義におけるメタと比較した場合、意味的な乖離が存在していると言える。例えば、巣や道具を用いる動物の例で考えると、どのようにして数ある行動の中から、道具や巣を作り、用いるプロセスの獲得に至ったかという部分こそがメタ的な行動の獲得であり、どのようにして道具をより効率的に使うかという部分は、メタ性が低い問題であると言える。本研究では、共有地の悲劇、造作物構築行動の獲得、いずれの問題においてもメタ性の低いメタ行動の獲得に対して焦点を当て研究を行い、進化的な獲得に成功している。本研究は、生き物の行動の内部評価自体を変えるような高次のメタ行動が発現する機構を考える上で重要な第一段階になり得るものであると考えている。実際の生物の生存競争下における進化において高次のメタ行動の獲得が見られるように、シミュレーションにおける高次のメタ行動の獲得でも、敵や餌となるものがある環境での進化による組み合わせ探索

を用いる手法は有用であると思われる。今後は、この延長線状での高次のメタ行動の獲得を扱う設計論へと結び付けたい。

### 7.3 今後の展望と予想される応用

メタ行動の導入による問題解決アプローチの応用例として、生物学的な意義と工学的な応用が考えられる。本研究は、自然の生物の行動の中にみられる間接的な行動に端を発している為、生物学的な意義との親和性が高いと考えられる。巣のような造作物構築行動の獲得だけではなく、設定の変更によって、異なるメタ行動に対して応用可能である。即ち、生物がどのようにして本来の目的とは異なる行動をメタ行動として獲得するに至ったのかという研究的な意義がある。

工学的な応用として、まず考えられる例は、自律的なロボットにおける行動獲得への応用である。造作物獲得の事例を適用する事で、評価関数で指定した機能を持つ造作物を自律的に構築するロボットを実現できる可能性がある。このようなロボットは、宇宙や深海など常に安定した通信が難しいような環境において特定のタスクを実行させる場合などに有効になると考えられる。また、造作物構築行動だけではなく、何らかの手順、プロセスの獲得に応用できる事を考えれば、プログラミングや何らかの操作手順の最適化などへ応用可能ではないかと考えられる。例えば、ある設計図の構築手順を提示する事で、最も構築プロセスの短い造作物設計方法を選択する手助けになると考えられる。その機能を用いれば、ある機能を持った思いがけない構造の提案をしつつ、人間による修正を加える事で、実用的かつ構築プロセスが短く、コストが低いものを提案可能になるようなソフトウェアなどが実現可能ではないだろうか。

以上のように、本研究で提案した手法は、生物学的な意義においても、工学的な応用においても、重要な役割を果たし得ると考えられる。



## 参考文献

- [1] Karl Sims (1994) Evolving virtual creatures. SIGGRAPH pp 15-22.
- [2] Karl Sims (1994) Evolving 3D morphology and behavior by competition. ALIFE 1(4): 353-372.
- [3] P. J. Funes, J. B. Pollack (1999) Computer evolution of buildable objects. Evolutionary Design by Computers, Morgan Kaufmann, San Francisco, pp. 387 – 403.
- [4] J. Rieffel, J. B. Pollack (2005) Evolving assembly plans for fully automated design and assembly. Proceedings of the 2005 NASA/DOD Conference on Evolvable Hardware, pp 165-170
- [5] J. Rieffel, J. B. Pollack (2005) Evolutionary fabrication: the emergence of novel assembly methods in artificial ontogenesis. SEEDS workshop at the 2005 GECCO, ACM Press, pp 265-272
- [6] J. H. Holland (1975) Adaptation in natural and artificial systems. MIT Press, Cambridge
- [7] Stefano Nolfi and Dario Floreano (2000) Evolutionary Robotics, The MIT Press
- [8] 吉倉眞 (1987)クモの生物学, 学会出版センター, ISBN 4-7622-0504-4
- [9] Russell Smith (2000-2008) Open dynamic Engine, URL <http://www.ode.org>
- [10] 出村公成 (2007) 簡単！実践！ロボットシミュレーション, 森北出版, ISBN 978-4-627-84691-3

- [11] Thomas Miconi and Alastair Channon (2006) An Improved System for Artificial Creatures Evolution, *Artificial Life 10*, The MIT Press, pp. 255-261
- [12] Nicolas Chaumont, Richard Egli and Christoph Adami (2006) Evolution of Virtual Catapults, *Artificial Life 10*, The MIT Press, pp. 262-267
- [13] Peter Krcak (2007), Evolving virtual creatures revisited, *GECCO'07*, pp341-350
- [14] Peter Krcak (2008), Towards Efficient Evolution of Morphology and Control, *GECCO'08*, pp. 287-288
- [15] Peter Krcak (2008), Towards Efficient Evolutionary Design of Autonomous Robots, *Lecture Notes in Computer Science Vol. 5216*, Springer, pp153-164
- [16] Kenneth O. Stanley and Risto Miikkulainen (2002) Evolving Neural Networks through Augmenting Topologies, *Evolutionary Computation 10(2)*:99-127, The MIT Press
- [17] Stefano Nolfi and D. Parisi (1995) Evolving non-trivial behaviors on real robots: an autonomous robot that pick up objects., In M.Gori and G.Soda (Eds.), *Topics in Artificial Intelligence. Proceedings of the 4th Congress of the Italian Association of Artificial Intelligence*. Berlin: Springer-Verlag, pp. 243-254
- [18] Stefano Nolfi (1997) Using emergent modularity to develop control system for mobile robots. *Adaptive Behavior (5)* 3-4:343-364

- [19] Stefano Nolfi (1997) Evolving non-trivial behaviors on real robots: a garbage collecting robot. *Robotics and Autonomous System*, 22: 187-198
- [20] Floreano D. & Nolfi S. (1997) Adaptive behavior in competing co-evolving species. In P. Husband & I. Harvey (Eds), *Proceedings of the Fourth Conference on Artificial Life*, MIT Press, Cambridge, MA, 378-387
- [21] Floreano D. & Nolfi S. (1997) God Save the red Queen! Competition in co-evolutionary robotics. In J.R.Koza, K.Deb, M.Dorigo, D.Foegel, B.Garzon, H.Iba & R.L.Riolo (Eds.) *Genetic Programming 1997: Proceedings of the Second Annual Conference*. San Francisco, CA: Morgan Kaufmann, pp. 398-406
- [22] Nolfi S. & Floreano D. (1998) Co-evolving predator and prey robots: Do 'arm races' arise in artificial evolution? *Artificial Life*, 4 (4), 311-335
- [23] Baldassarre G., Nolfi S., Parisi D. (2003). Evolution of collective behavior in a team of physically linked robots. In S. Cagnoni et al. (eds.) *Applications in Evolutionary Computing*. Lecture Notes in Computer Science, Berlin: Springer Verlag, pp. 581-592.
- [24] Mondada F., Pettinaro G., Guigrard A., Kwee I., Floreano D., Deneubourg J-L, Nolfi S., Gambardella L.M., Dorigo M. (2004). Swarm-bot: A new distributed robotic concept. *Autonomous Robots*. vol. 17 (2-3), pp. 193-221.

- [25] Baldassarre G., Parisi D., Nolfi S. (2004). Coordination and behavior integration in cooperating simulated robots. In S. Schaal, A. Ijspeert, A. Billard, S. Vijayakumar, J. Hallam and J-A. Meyer (Eds.), *From Animals to Animats 8: Proceedings of the VIII International Conference on Simulation of Adaptive Behavior*, pp. 385-394. MIT Press, Cambridge, MA.
- [26] Baldassarre G., Parisi D., Nolfi S. (2006). Distributed Coordination of Simulated Robots Based on Self-Organisation, *Artificial Life*. 12(3):289-311.
- [27] Baldassarre G., Trianni V., Bonani M., Mondada F., Dorigo M. & Nolfi S. (2007). Self-organised coordinated motion in groups of physically connected robots. *IEEE Transactions on Systems, Man, and Cybernetics*, 37(1):224-239.
- [28] Baldassarre, G., Nolfi, S. (2009). Strengths and synergies of evolved and designed controllers: a study within collective robotics. *Journal of Artificial Intelligence*. Vol. 173, pp. 857-875,
- [29] Thimo Krink and Fritz Vollrath (1997) Analysing Spider Web-building Behavior with Rule-based Simulations and Genetic Algorithms, *Journal of Theoretical Biology* Vol. 185, pp. 321-331

- [30] Garret Hardin (1968) The Tragedy of the Commons, *Science*(162), pp.1243-1248
- [31] R.M. Dawes (1980) Social Dilemmas, *Annu. Rev. Psychol.* 31, 169-193.
- [32] Keiji Suzuki (2002) Coevolving Levy Plans and Payoff Predictions in the Tragedy of the Commons, *Agent-based Approaches in Economic and Social Complex Systems*, IOS press, pp. 122-133.
- [33] 山下 倫央, 鈴木 恵二, 大内 東 (2000) Iterated Multiple Lake Game における社会的ジレンマに対するプレイヤー群の挙動に関する考察, *計測自動制御学会論文集*, Vol. 36, No. 2, pp. 195-203.
- [34] 盛山 和夫、海野 道郎, (1991) 秩序問題と社会的ジレンマ”, *ハーベスト社*, pp. 137-166.
- [35] 鈴木 光男 (1994) 新ゲーム理論, *勁草書房*, pp. 18-22 等.
- [36] 宮西 啓司 (2005) 修士論文: メタエージェント化機能に基づく社会シミュレーション
- [37] Rodney Brooks (1986 ) A robust layered control system for a mobile robot, *Robotics and Automation Volume 2 Issue 1*, pp. 14-23

[38] Rodney Brooks (1991) Intelligence Without Representation, Artificial Intelligence

No.47 pp. 139-159

## 業績一覧

### 査読付き論文

- [1] Kenta Oomiya, Keiji Miyanishi, Keiji Suzuki  
Balancing Sociality in Meta-agent Approach  
Autonomous Robots and Agents, Springer, 2007, pp. 235-239
  
- [2] Kenta Oomiya, Keisuke Nakamura, Keiji Suzuki  
Evolutionary acquisition of behavior to build structural objects by virtual creatures  
Artificial Life and Robotics, Springer, 2009, (採択済み)

### 国際発表

- [1] Kenta Oomiya, Keiji Miyanishi, Keiji Suzuki  
Design of Tax Charging Players in the Tragedy of the Common  
Intelligent and Evolutionary System 2005 (Hakodate), CD-ROM,  
2005/11/19-20
  
- [2] Kenta Oomiya, Keiji Miyanishi, Keiji Suzuki  
Evolutionary Effects of Meta-agent Approach in the Tragedy of the Commons  
Artificial Computational Economics & Social Simulation 2005 (Hakodate),  
CD-ROM, 2005/12/17-18

- [3] Kenta Oomiya, Keiji Miyanishi, Keiji Suzuki  
Weighted Evaluation of Meta-agent Approach in the Tragedy of the  
Common  
SCIS & ISIS 2006 (Tokyo) Paper No. SU-B2-1 (CD-ROM), 24th,  
September, 2006
- [4] Kenta Oomiya, Keiji Miyanishi, Keiji Suzuki  
Balancing Sociality in Meta-agent Approach  
3<sup>rd</sup> International Conference on Autonomous Robots and Agents( ICARA  
2006)  
Palmerston North, New Zealand,  
pp.605-610, 12-14 December 2006
- [5] Kenta Oomiya, Keisuke Nakamura, Keiji Suzuki  
Evolutionary Acquisition of Behaviors Building Structural Objects by  
Virtual Creatures  
International Symposium on Artificial Life and Robotics 14<sup>th</sup>,  
2009/02/05-07
- [6] Kenta Oomiya, Keiji Suzuki  
Influences of Mid-term Evaluations on Acquisition by Virtual Creatures  
COMPLEX' 09, 2009/11/4-7 in Tokyo



国内学会・シンポジウム等における口頭発表

[1] 大宮健太, 宮西啓司, 鈴木恵二

共有地の悲劇におけるメタエージェント化機能の進化的な効果

第 143 回 知能と複雑系研究会 2006 年 3 月 28 日

[2] 大宮健太, 鈴木恵二

負荷分散問題への社会性エージェントアプローチ

2006 年度精密工学会北海道支部学術講演会, pp 9-10 2006 年 9 月 2 日

[3] 大宮健太, 鈴木恵二

グリッドコンピューティング向けエージェントベースアプローチ

情報処理北海道シンポジウム 2006, pp 54-55, 講演番号 B-8 2006 年 10 月 12 日

[4] 大宮健太, 中村圭介, 鈴木恵二

仮想生物による造作物構築行動の進化的獲得に関する基礎研究

電気学会 2008 年度電子情報システム部門大会, 2008 年 8 月 20 日～22 日

[5] 大宮健太, 中村圭介, 鈴木恵二

造作物構築行動の進化的獲得を行う仮想生物に関する基礎研究

情報処理北海道シンポジウム 2008, ポスター番号 A-10, 2008 年 9 月 19 日

[6] 大宮健太、中村圭介、鈴木恵二

仮想生物による身体的な特徴を用いた造作物構築行動

第18回インテリジェント・システム・シンポジウム(FAN2008)、2008年10月23、24日

国内学会・シンポジウムなどにおけるポスター発表

[1] 大宮健太、中村圭介、鈴木恵二

仮想生物による進化的な造作物の生成に関する基礎研究

Robomec 2008、2008年6月5日～7日

## 謝辞

本研究の過程において、終始懇切なるご指導とご鞭撻を賜り、本論文をまとめるに際してご助言を頂いた、公立はこだて未来大学システム情報科学部三上貞芳教授、北海道大学大学院情報科学研究科鈴木恵二教授に、心より感謝いたします。また、本論文の審査過程において、数々のご助言とご指導を頂きました、大沢教授、松原教授、ハルトノ准教授に深謝申し上げます。

これまでの研究生生活において、多大なご指導、ご鞭撻を多くの方々から賜り、支えて頂き、こうして博士論文をまとめる事が出来ました。また、留学先である Dalhousie 大学にて NeuroEvolution 手法に関する指導を担当して頂いた、Professor Dirk Arnold に深謝申し上げます。英語に関するアドバイスを頂いた沼田先生と Connection's Café の講師陣の方々、その他ご助言を頂いた方々、研究室の皆様、事務の方々に深く感謝致します。また、資金面で御援助頂いた日本学生支援機構、道南学術支援機構の皆様方にも感謝致します。